

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 12, 2005, 10:30:41 ; Search time 10000 Seconds
(without alignments)
5194.399 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 5580

Sequence: 1 MAPTGVLSLLLVITAGC.....VSMNGSIRNGASFVCSKDR 1072

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0		0.5	
Ygapop 6.0		0.5	7.0
Delop 6.0		7.0	7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO spool_p/US10764390/runat_12102005_110203_22329/app_query.fasta_1.1223
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOPC=0.1 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10764390@cgn 1 1 6626 @runat 12102005_110203_22329 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sv.*

13: gb.un.*

14: gb.vi.*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5574	99.9	6791	9 AB002317	AB002317 Homo sapi
2	4255.5	76.3	4971	10 AX122246	AX122246 Mus muscu
3	4092	73.3	2433	6 CQ726592	CQ726592 Sequence
4	2435.5	43.6	4379	10 BC028869	BC028869 Mus muscu

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	2405	43.1	4183	9	BC014530
6	2404	43.1	3908	6	AX195570
7	2404	43.1	4000	9	AY163234
8	2385	42.7	4001	9	AF275679
9	2269	40.7	2970	6	CQ727300
10	2184.5	39.1	2796	6	BD160555
11	2184.5	39.1	2796	6	AX883831
12	2184.5	39.1	2796	9	AK024287
13	2122.5	38.0	2743	9	BC031672
14	2005.5	35.9	4423	9	AB058740
15	1817.5	32.6	2501	6	BD057820
16	1798	32.6	5555	6	BD183361
17	1675.5	30.0	2811	10	BC022154
18	1451.5	26.0	3345	6	CQ589125
19	1443.5	25.9	2223	6	BD157975
20	1443.5	25.9	2223	6	AX879698
21	1443.5	25.9	2223	9	AK022734
22	1426	25.6	3727	3	BT010098
23	1395	25.0	2322	9	AK090878
24	1344	24.1	28692	9	AL512385
25	1280.5	22.9	5775	6	CQ589124
26	1280.5	22.9	94236	2	AC020539
27	1280.5	22.9	185790	3	AC091221
28	1280.5	22.9	185755	3	AC010037
29	1280.5	22.9	281917	3	AE003556
30	1274.5	22.8	3133	3	AK114825
31	1267.5	22.7	3121	3	AK113004
32	1166	20.9	57708	2	AC130340
33	996	17.8	1000	5	CR353665
34	931.5	16.7	1874	9	BC007645
35	886.5	15.9	256608	10	AL589699
36	880.5	15.8	245394	2	AC125756
37	845	15.1	772	6	BD148725
38	845	15.1	772	6	AX868663
39	823	14.7	818	6	AX677257
40	786	14.1	57349	2	AC135073
41	734	13.2	720	5	EX933224
42	681	12.2	580	6	BD071994
43	640	11.5	578	6	BD071762
44	525.5	9.4	67498	2	AC012648
45	506.5	9.1	46913	2	AC023733

ALIGNMENTS

RESULT 1	AB002317	Homo sapiens mRNA for KIAA0319 gene, partial cds.	6791 bp	mRNA	linear	PRI 10-JAN-2004
LOCUS	AB002317	Homo sapiens mRNA for KIAA0319 gene, partial cds.				
DEFINITION	AB002317	Homo sapiens mRNA for KIAA0319 gene, partial cds.				
ACCESSION	AB002317	Homo sapiens mRNA for KIAA0319 gene, partial cds.				
VERSION	AB002317.1	GI:2224578				
KEYWORDS						
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS		1 Nagase, T., Ishikawa, K., Nakajima, D., Ohira, M., Seki, N., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.				
TITLE		Prediction of the coding sequences of unidentified human genes.				
		VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro				
JOURNAL		DNA Res. 4 (2), 141-150 (1997)				
MEDLINE		97349984				
PUBMED		9205841				
REFERENCE		2 (bases 1 to 6791)				
AUTHORS		Ohara, O., Nagase, T., Kikuno, R. and Nomura, N.				
TITLE		Direct Submission				
JOURNAL		Submitted (28-MAR-1997) Osamu Ohara, Kazusa DNA Research Institute; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan				
FEATURES		(E-mail:cdnaif@kazusa.or.jp, Tel: +81-438-52-3913)				
source		Location/Qualifiers				
		1. .6791				

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HG0378"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
1. .6791
/gene="KIAA0319"
<401..3730
/gene="KIAA0319"
/notes="Start codon is not identified"
/codon_start=1
/protein_id="BAA20777.2"
/db_xref="GI:40788212"
/translation="POPLRNLPAPKKHOKRLAPNSIQGLRSLPSPTVWCTWAPPTGV
LSGLLLVVIAGCARQKQCEGTYNSNAVISPNEITRINRVSHTPPVVDCTAACCDLS
SCDLAWFGRCLVSCPKHCEPKMGPIRSYLFVLVRPVRQAPLQDQKQYQKQSGSAR
GSPGTWGDSPEDIRKDPFLGKDWGLEMSYADDYRELEKDLQPSGKQPGRSAR
YTDWGLPGSEGAFFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLP
TTPSGEVLKKAQOLQSSNSGKVLPSHSLPPASLELSSVTVKEKSPVLVTPP
GSTEHSIPTTPSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLILTLDPNEVELK
AFVAPAPPVTITINWNLISHPTDQGEIKQGHKQNLNLSQLSVGLVFKVTVSSEN
AFGEFVNTVTPARVNLPPVAVVSPQIQELTLPLTSALIDGSDSTDDTEIVSHWE
ENGPFIEKTSVDSVPLNSLNDPCNRYFRLTVDSDGATNSTTAAITVNNADVPP
VANAGPNHTITLTPQNSITLNGQSSDDHQIVLYEWSLPGSGKHVMOGVOTPYLHL
SAMOEGDITFOLKVTDSRQOSTAVVTIVOPENNRPPVAVAGPKELIFPVESATLD
GSSSDDHGIVFVHWEHVRGSPAVEMENIDKALATVTVGLVGTYPHRLTVKDDQGLSS
TSLTVAVKKNNSPPRARAGGHVLLVFNNSITLDGSRSTDQRIIVSYLWIRDOQSP
AAGDVIIDGVALQTLNLBEGVYTFHLRVDSQASDTDTATVEVDPDRKSGLVE
LTLQGVGQUTQEQKDTLRQALVLLNVLDSIKVKQIRAHSDLSLTVIVYVQSRPPF
KVLKAAEARNLHMRLSKGKDLFLFKVLVDTAGCLLKCSGHGHCDDPLTKRCICSHL
WENLILQRYIWDGESNCEWSIFVTVLAFTLVLGTGFTWLCICCCCKRQKIRKKT
KVTILDNDEQERMEHLRPKYGIKHRSTEHNSLWSESEFSDQDTIFSRKMERGNP
KVSMMNGSIRNGASFSYCSKDR"

Alignment Scores:
Pred. No.: 3 04e-291 Length: 6791
Score: 5574.00 Matches: 1069
Percent Similarity: 99.72% Conservativeness: 0
Best Local Similarity: 99.72% Mismatches: 3
Query Match: 99.89% Indels: 0
DB: 9 Gaps: 0

US-10-764-390-3 (1-1072) x AB002317 (1-6791)

QY 1 MetAlaProThrGlyValLeuSerSerLeuLeuLeuValThrIleAlaGlyCys 20
|
|
|
Db 512 ATGGCGCCCCACAGGTGTGCTCTTCAATGCTGCTGCTGGTGACAAATGCGAGTTGT 571
|
|
|
QY 21 AlaArgGlyGlnCysSerGluGlyArgThrTyrSerAsnAlaValIleSerProAsnLeu 40
|
|
|
Db 572 GCCCGTAAGCAGTGCACGAGGGGAGACATATCCAAATGCAATATTCACCTAATCTG 631
|
|
|
QY 41 GluThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAlaAla 60
|
|
|
Db 632 GAACACACAGATCATCGGGGTGCTCACACCTTCCCTGTCTGAGACTGCACGGCGGCT 691
|
|
|
QY 61 CysCysAspLeuSerSerCysAspLeuAlaTrpTrpPheGluGlyArgCysTyrLeuVal 80
|
|
|
Db 692 TGCTGTGACCTGTCAGCTGACCTGGCTGGCTGGTTCGAGGGCGGCTGCTACCTGGTG 751
|
|
|
QY 81 SerCysProHisGlyAsnGlnCysGluProLysIleMetGlyProIleArgSerTyrLeu 100
|
|
|
Db 752 AGCTGCCCCACAAAGAGAACTGTGAGCCCAAGAGATGGGCCCAATCAGGCTTATCTC 811
|
|
|
QY 101 ThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMet 120
|
|
|
Db 812 ACTTTTGTGCTCCGGCTGTTCAGAGGCTTGCACAGCTGCTGGACTATGGGACATGATG 871
|
|
|
QY 121 LeuAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAspIleArgLysAsp 140
|
|
|

```

ORIGIN

```

Db 872 CTGAACAGGGGCTCCCTCGGGGATCTGGGGGAGCTACCTCGAGGATATCAGAAAGGAC 931
QY 141 Leu***PheLeuGlyLysAspTTPGlyLeuGluMetSerGluTyr***AspAspTyr 160
|
|
|
Db 932 TTGCCCTTTTAGGCAAGATTTGGGCTTAGAGGAGATGTCTGAGTAGCGAGATGACTAC 991
|
|
|
QY 161 ArgGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAla 180
|
|
|
Db 992 CGGGAGCTGGAGAGGACCTCTTGCAACCCAGTGGCAAGCAGGAGCCCGAGAGGAGTGCC 1051
|
|
|
QY 181 GluTyrThrAspTTPGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGly 200
|
|
|
Db 1052 GAGTACACGAGCTGGGGCTTCTGCCGGGACGAGCGGGGGCTTCAACTCTCTCTGTGGA 1111
|
|
|
QY 201 AspSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAsnGlu 220
|
|
|
Db 1112 GACAGTCTCGGTGCGAGCGGAGCGAGCAGGAGCCCTGAGCTCCATTACCTGATGAG 1171
|
|
|
QY 221 SerAlaSerThrProAlaProLysLeuProGluArgSerValLeuLeuProLeuProThr 240
|
|
|
Db 1172 TCGGCTTCAACCCCTGCCCCAAAACCTCCCTGAGAGAGTGTGTGTTGCTTCCCTTGC 1231
|
|
|
QY 241 ThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGlnGlnSer 260
|
|
|
Db 1232 ACTCCATCTTCAGGAGAGGTGTGGAGAAAGAAAAGGCTTCTCAGCTCCAGGAACAATCC 1291
|
|
|
QY 261 SerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProProAlaSerLeu 280
|
|
|
Db 1292 AGCAACAGCTCTGGAAAAGAGGTTCTAATGCTTCCCATAGTCTTCTCCGCGCAGCCTG 1351
|
|
|
QY 281 GluLeuSerSerValThrValGluLysSerProValLeuLeuThrValThrProGlySerThr 300
|
|
|
Db 1352 GAGCTCAGCTCAGTCACCGTGAGAAAAAGCCAGTGCTCACAGTCACCCCGGGAGTACA 1411
|
|
|
QY 301 GluHisSerIleProThrProThrSerAlaAlaProSerGluSerThrProSerGlu 320
|
|
|
Db 1412 GAGCACAGACTCCCAACACCTCCCACTAGCGCAGCCCTCTCGAGTCCACCCCATCTGAG 1471
|
|
|
QY 321 LeuProIleSerProThrThrAlaProArgThrValLysGluLeuThrValSerAlaGly 340
|
|
|
Db 1472 CTACCCATATCTCTTACCACCTGCTCCAGGACAGTGAAGAAGAACTTACGGTATCGGCTGA 1531
|
|
|
QY 341 AspAsnLeuIleIleThrLeuProAspAsnGluValGluLeuLysAlaPheValAlaPro 360
|
|
|
Db 1532 GATAACCTAAATTATAACTTTTACCCGACAAATGAAGTTGAACCTGAAGGCCCTTTGTCGCCA 1591
|
|
|
QY 361 AlaProProValGluThrThrTyrAsnTyrGluTrpAsnLeuLeuIleSerHisProThrAsp 380
|
|
|
Db 1592 GCGCCACCTGTAGAAAACAACCTTACAACTATGAATGGAAATTTTAATGAAGCCACCCACAG 1651
|
|
|
QY 381 TyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerVal 400
|
|
|
Db 1652 TACCAAGGTGAATTAATAACAAGACACAAAGCAAACTCTTAACCTCTCTCAATGTCCGTC 1711
|
|
|
QY 401 GlyLeuTyrValPheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPheVal 420
|
|
|
Db 1712 GGACTTTATGTCTTCAAAAGTCACCTGTTCTAGTGAAGAACCGCTTTGGAGAGAGATTTGTC 1771
|
|
|
QY 421 AsnValThrValLysProAlaArgValGlnLeuProProValAlaValValSerPro 440
|
|
|
Db 1772 AATGTCACTGTTAAGCCTGCCAAGAGAGTCAACCTGCCACCTGTGAGCAGTTGTTCTCCC 1831
|
|
|
QY 441 GlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThr 460
|
|
|
Db 1832 CAACTGCAAGAGCTCACTTGGCTTTGAGTCAAGCCCTCATTTGATGGGAGCCCAAGATACA 1891
|
|
|
QY 461 AspAspThrGluIleValSerTyrHisTrpGluGluIleAsnGlyProPheIleGluGlu 480
|
|
|
Db 1892 GATGATACTGAAATAGTGAGTTATCATTTGGGAAGAAATAAAGCGGCCCTTTCATAGAAGAG 1951
|
|
|
QY 481 LysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSer 500
|
|
|
Db 1952 AAGACTTCAAGTGAAGTCTCCCGCTTACGCTTGTCTAACTTGTATGCTCTGTAACATATAGT 2011
|
|
|

```

Qy	501	PheArgLeuThrValThrAspSerAspGlyValAlaThrAsnSerThrThrAlaAlaLeuIle	520
Db	2012	TTCAAGTGTGACTGTTACAGACTCGGACGGAGCCATAACTCTACTCAACTGTCAGAGCCCTAATA	2071
Qy	521	ValAsnAsnAlaValAspTyrProProValAlaAsnAlaGlyProAsnHisThrIleThr	540
Db	2072	GTGAACAATCTGTGGACTACCCACCACTGGTCTAATCGAGACCAATACACCATCAACT	2131
Qy	541	LeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleVal	560
Db	2132	TTGCCCCAAAACTCCATCACTTTGAAATGGAAACCAGAGCAGTCAGCATCACAGATTGTC	2191
Qy	561	LeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHisValValMetGlnGlyVal	580
Db	2192	CTCTATGAGTGGTCCCTGGGTCTCTGGGAGTGGGGCAAACTGTGGTTCATCGAGGGAGTA	2251
Qy	581	GlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAspTyrThrPheGlnLeuLys	600
Db	2252	CAGACGCCATACCTTCATTTATCTGCNAATCGAGGAGGAGATTATACATTTTCAGCTGAAG	2311
Qy	601	ValThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnProGlu	620
Db	2312	GTGACAGATTCTTCAAGGCCAACGTCTACTGCTGTAGTGACTGTGATGTGTCAGAGCTGAA	2371
Qy	621	AsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProValGlu	640
Db	2372	AACAATAGACCTTCAGTGGCTGTGGCGCGCCCTGATAAAGAGCTGATCTTCCAGATGGAA	2431
Qy	641	SerAlaThrLeuAspGlySerSerSerSerAspAspHisGlyIleValPheTyrHisTrp	660
Db	2432	AGTGCTACCTTGGATGGAGCAGCAGCAGCGATGACCAACGGCATTTGTTCTTACCACCTGG	2491
Qy	661	GluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAlaThr	680
Db	2492	GAGCACGTCAGAGCCCCAGTCGAGTCGAGATGGAAATATTGACAAAGCAATAGCCACT	2551
Qy	681	ValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGly	700
Db	2552	GTCACTGGTCTCCAGTGGGGAGCTACCTCTTCGTTTGACAGTGAAGACCAACGACGGGA	2611
Qy	701	LeuSerSerThrSerThrLeuThrValAlaValLysLysGluAsnAsnSerProProArg	720
Db	2612	CTGAGCAGCAGCTCCACCTCACTGTGGCTGTGAGAGAGGAAATATAGTCTCTCCAGA	2671
Qy	721	AlaArgAlaGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAspGly	740
Db	2672	GCCCGGCTGTGTGCAGACATGTTCTTGTGCTTCCCAATAATTCCATTACTTTGGATGGT	2731
Qy	741	SerArgSerThrAspAspGlnArgIleValSerTyrLeuTrpIleArgAspGlyGlnSer	760
Db	2732	TCAAGGTCTACTGATGACCAAGAAATTTGTGTCTTATCTGTGGATCTCGGGATGGCCAGGT	2791
Qy	761	ProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsn	780
Db	2792	CCAGCAGCTGGAGATGTCATCGATGGCTCTGACCACAGTGTGGCTCTGCAGCTTACGAAT	2851
Qy	781	LeuValGluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAsp	800
Db	2852	CTGTGGAGGGGTGTACACTTTTCCACTTGCAGTCCCGACAGTCAGGGGGGCTCGGAC	2911
Qy	801	ThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValGluLeu	820
Db	2912	ACAGCACTGCCACTGTGGAAAGTGCACCAACCTAGAAAGATGGCTGTGGTGGAGCTG	2971
Qy	821	ThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuValArgGln	840
Db	2972	ACCTCGAGGTTGTTGTTGGSCAGCTGCACAGAGCAGCGGAGGACACCTTGTGAGGCAG	3031
Qy	841	LeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLysIleArgAlaHis	860
Db	3032	CTGGCTGTGCTGTGAACGTCGTGGACTCGACATTAAGGTCCAAGATTCGAGAGATTCGGGCCAC	3091

Qy	861	SerAspLeuSerThrValIleValPheTyValGlnSerArgProPropheLysValLeu	880
Db	3092	TCGGATCTCAGCACCGGTGATGTGTGTTATGTATCAGAGCAGCGCGCTTTTCAAGGTTCTC	3151
Qy	881	LysAlaalaGluValAlaIleArgAsnLeuHisMetArgLeuSerLysGluLysAlaAspPhe	900
Db	3152	AAAGCTGCTGAGTGGCCCGAATCTGCACATCGGGCTCTCAAGAGAGAAGGCTGACTTC	3211
Qy	901	LeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGlyHis	920
Db	3212	TTGCTTTTCAAGGTTCTTGGGGTGTATACAGCAGGTTGCCTTCTGAAGTGTCTCTGGCCAT	3271
Qy	921	GlyHisCysAspProLeuThrLysArgCysIleCysSerHisIleuTrpMetGluAsnLeu	940
Db	3272	GGTCACCTGCGACCCCTCTACAAAGCGCTGCATTTGCTCTCACATTATGGATGGAGAACCCT	3331
Qy	941	IleGlnArgTyrlleTrpAspGlyGluSerAsnCysGluTrpSerIlePheTyValThr	960
Db	3332	ATACAGCGTTATATCTGGGATGGAGAGAGCACTGTGAGTGGAGATATATTCTATGTGACA	3391
Qy	961	ValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCysCys	980
Db	3392	GTGTGGCTTTTACTCTTATTGTGCTAAACAGGAGGTTTCACTTGGCTTTGCATCTGTCTGC	3451
Qy	981	CysLysArgGlnLysArgThrLysIleArgLysLysThrLysTyThrIleLeuAspAsn	1000
Db	3452	TGCAAAAGACAAAAAGAGACTAAATCAGGAAAAACAAAGTACACCATCTCTGGATTAAC	3511
Qy	1001	MetAspGluGlnGluArgMetGluLeuArgProLysTyrcGlyIleLysHisArgSerThr	1020
Db	3512	ATGATGAACAGGAAGAAATGGAACTGAGGCCCAATATGGTATCAAGCACCGAAGCACA	3571
Qy	1021	GluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIle	1040
Db	3572	GAGCACAACTCCAGCCTGATGGTATCCGAGTCTGAGTTTCACAGTGACCCAGGACACAATC	3631
Qy	1041	PheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerIleArg	1060
Db	3632	TTACGCCGAGAAAGATGGAGAGAGGGAATCCAAAGGTTTCCATGAATGGTTCCATCAGA	3691
Qy	1061	AsnGlyAlaSerPheSerTyrcSerLysAspArg	1072
Db	3692	AATGGAGCTTCCTTCAGTTATTCCTCAAGAGGACAGA	3727
RESULT 2			
AKI22246		4971 bp mRNA linear ROD 15-MAR-2003	
LOCUS	AKI22246	Mus musculus mRNA for mKIAA0319 protein.	
DEFINITION	AKI22246		
ACCESSION	AKI22246		
VERSION	AKI22246.1	GI:28972148	
KEYWORDS	FLI CDNA.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

3'-end one pass sequencing.			
FEATURES	Location/Qualifiers		
source	1..4971		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/clone="mbg09941"		
	/tissue_type="brain"		
	/dev_stage="adult"		
	/note="vector:modified pBC SK+"		
gene	1..4971		
	/gene="mKIAA0319"		
CDS	<249..3500		
	/note="mKIAA0319"		
	/note="CDS is predicted by in silico analysis. Start codon is not identified."		
	/evidence=not_experimental		
	/product="mKIAA0319 protein"		
	/protein_id="BAC65528.1"		
	/db_xref="GI:28972149"		
	/translation="GRMVSPRGVLSLLLLAAMAGSSQCSGRYSDALISNPETIRIMVSTFSGDCTAACCDLLTCDLAWFEGSCVLYKMRSENCEPTTGPIRSYLTVRRVPQGLDDYGMWLRSGSPGAWGSDLELDKDLPLGKGGPBEETISDYKDLERGLQSPNQDPRGSABYPDWISLLPSNCGFNATATGDSAAKMEKLDPTPHLDQRLQALANESTWSPTPGHSSISSVWFPSSASPLPTEGLEGETLQLOQPSSNSGKEVPWPSHNPSPASLESSPATTEKNSNFTVTPRSKHSSTPFTSTVLGTLPDPPWP LPTASRTWKALAVSAGDNILVLTLPDREAELKASVEPAPADTVSYEWSLMSHPDVFQGIKQENKPLTHLSQSLVGLYAFRAVAVSENAFGBGYVNTVMPAARVNQPPVAVVSPOQLSESLTSLIDGSDTDDTEIVSYHWEVDGFFLGEFFPADTILRLSNLVPVNTYFRLITDSGATNSTASLWIRDAVDYPPVANAGPNTITLQNTIILNQSSDDHOIVLVYFWFAGGGSKEMWQGAOTPYLHLSQEGEYTPQLMWTSSGQSQSTALVAVTQAEENQAPVAVAGPDKELFPVQSATLDGSRSSDDHGIVCVYHWHIRGPSAVEMENVDKAIATVTGLQVGIYFRLTVRQGLSSTSLTVAYKKENSPPPRAQAGRHVLILPNSITLDTGSDRDRGIVSYLWIRDGSPGAGDVITGSDHRAALQTLNLVEGYITFHLVTSQGSADSDAATVEVLPDPKKGVLVLIQVGEQLTEQOKETFLVRQLAVLLNVLDSVTKIQAHTDVTSTVIVFYQSGSPFKFLRAAAVARNLHKRLSKKEAEFLLFKVLRVDTGAVSLKSGHGHCDPTKRCICSQLMWNLIQRYWMDGSENCEWVFVVAALALTLLTAGVNSLWLCICCCRRKRTKIRKTKYTLTDSMDQERWELRPKYGIKHSRSTENSLMVSSEFSDQDTLFSRERMERGVLKSLNGCARNGVSPGYYSKDR"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,12e-220	Length:	4971
Score:	4255.50	Matches:	820
Percent Similarity:	84.20%	Conservative:	91
Best Local Similarity:	75.79%	Mismatches:	160
Query Match:	76.26%	Indels:	11
DB:	10	Gaps:	4
US-10-764-390-3 (1-1072) x AK122246 (1-4971)			
Qy	1	MetAlaProThrGlyValLeuSerSerLeuLeuLeuValThrLeAlaGlyCys	20
Db	255	ATGTGTGCCACACGAGTACTGCTTCATTGCTGCTGGCAGCCATGGCAGGAGGC	314
Qy	21	AlaArgLysGlnCysSerGluGlyArgGlyThrSerAsnAlaValIleSerProAsnLeu	40
Db	315	AGTTCTCAGCAGTGCTCTGAGGGCAGGACTTACTCCGATGCCATCTTCCACCTAACCCG	374
Qy	41	GluThrThrArgIleMetArgValSerHisThrPheProValValAlaAspCysThrAlaAla	60
Db	375	GAACCATCAGATCATGCGGGTGCTCAAACCTTCTCCGTGGGAGACTGCACGGCGCT	434
Qy	61	CysCysAspLeuSerSerCysAspLeuAlaTrpTrpPheGluGlyArgCysTyrLeuVal	80
Db	435	TGCTGTGACTTGCTCACCTGTGACCTGGCTGGTGGTTTTCAGGGCAGCTGCTATCTGGTG	494
Qy	81	SerCysProHisLysGluAsnGlyGluProLysLysMetGlyProIleArgSerTyrLeu	100
Db	495	AAATGCAATGGCTCGGAGAAATTGGACGCCACGACACAGGCCCATCCGGTCTTACCTC	554
Qy	101	ThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMet	120
Db	555	ACTTTCGTGCGCAGACCTGTCCAGAGGCCCGGCGAGCTGTGGACTATGGAGACATGATG	614
Qy	121	LeuAsnArgGlySerProSerGlyTrpGlyAspSerProGluAspIleArgLysAsp	140
Db	615	CTGAGCAGGGCTCCCTCCCTCAGGAGCTTGGGGAGACTCCCTTGGAGACTCAGGAAGAC	674
Qy	141	Leu***PheLeuGlyLysAspTrpGlyLeuGluMetSerGluTyr***AspAspTyr	160
Db	675	TTGCCCTTCTTGGCAAGACGCGGGACACGAGGAGACCACTGAGTACTCAGATGAATAC	734
Qy	161	ArgGluLeuGluLysAspLeuGlnProSerGlyLysGlnGluProArgGlySerAla	180
Db	735	AAAGACCTGGAGCGGGCTCTCTGCAGCCCTCCAAACAGCAGAACCTTAGAGGTAGCGCA	794
Qy	181	GluTyrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsn---SerSerVal	199
Db	795	GAGTACCCCGACTGGAGCTCTGCTGCCAGACAGGAGGAGGTTTCAATGCTACAGCTACA	854
Qy	200	GlyAspSerProAlaValProAlaGluThrGlnGlnAspPro-----	213
Db	855	GGAGACAACCTCAGCTGCTTCCATGGAGAAGCTGCAGGATCCACGCCCCACCCACTGGAC	914
Qy	214	-----GluLeuHisTyrLeuAsnGluSerAlaSerThrProAlaProLysLeuProGlu	231
Db	915	CAGGACGAGCTGCAGGCCCTGAAATGAGTCGACTTGTGTCCTTACACCCGGGCACCTCTTCA	974
Qy	232	ArgSerValLeuLeuProLeuProThrPro---SerSerGlyGluValLeuGluLys	250
Db	975	ATAACGAGTGTGGCCCTCTCTGCGAGCCCATTTACCTACAGAGGAGGACTGGAAGA	1034
Qy	251	GluLysAlaSerGlnLeuGlnSerSerAsnSerSerGlyLysGluValLeuMet	270
Db	1035	GAAGAGACTTTACAGCTCCAAGACAACCGAGCAATAGCTCTGGAAGAGAGGTCCAATG	1094
Qy	271	ProSerHisSerLeuProProAlaSerLeuGluLeuSerSerValThrValGluLysSer	290
Db	1095	CTTCCCATATAATCTCTCCCTGCCAGCCTGGAGTCTAGCCACGACGACGCAAAAAAAC	1154
Qy	291	ProValLeuThrValThrProGlySerThrGluHisSerIleProThrProProThrSer	310
Db	1155	TCCAACTTTACAGTCACGCCGAGGAGAGAAACACGACGCCCACTTTCTTACCAGT	1214
Qy	311	AlaAlaProSerGluSerThrProSerGluLeuProIleSerProThrThrAlaProArg	330
Db	1215	ACAGTCTCTCAGTGGGCTAACCCCTCTCCGTGGCCCTTGTCACTT---ACTGTTCCAGG	1271
Qy	331	ThrValLysGluLeuThrValSerAlaGlyAspAsnLeuIleIleThrLeuProAspAsn	350
Db	1272	ACAGTAAAGGCGCTTGCTGTCTGCTGGAGATAACCTAGTACTTAACCTTACCGGACAGA	1331
Qy	351	GluValGluLeuLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTyr	370
Db	1332	GAAGCAGAACTGAAGGCTCTGTTGAACGAGCGCCCTTCGAGATACAACTTACTCCTAT	1391
Qy	371	GluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGluIleLysGlnGlyHisLys	390
Db	1392	GAATGGAGTTTAATGAGCCACCAGTAGACTTCCAAGGTAAATCAACAAAGAAAAACAG	1451
Qy	391	GlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSer	410
Db	1452	CCGACTCTTCACTCTCAATTAATCTGTGGGACTCTATGCTTCCAGAGTGGCTGTCTTCT	1511
Qy	411	SerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgVal	430
Db	1512	AGTGAGAAATGCATTTGGAGAGGCTATGTCAATGTCACGGTTATGCCAGCTCAAGAGTC	1571
Qy	431	AsnLeuProProValAlaValSerProGlnLeuGlnLeuThrLeuProLeuThr	450
Db	1572	ACCAGGCCACCTGTAGTGTGTTCTTCTCCACAGACACAGAGAGCTCAGTGTGCTTTGACC	1631
Qy	451	SerAlaLeuLysAspGlySerGlnSerThrAspThrGluIleValSerTyrHisTrp	470

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 8.54e-212 Length: 2433
Score: 4092.00 Matches: 807
Percent Similarity: 80.34% Conservative: 2
Best Local Similarity: 80.14% Mismatches: 2
Query Match: 73.33% Indels: 196
DB: 6 Gaps: 5

US-10-764-390-3 (1-1072) x CQ726592 (1-2433)

QY 46 MetArgValSerHisThrPheProValValAspCysThrAlaAlaCysCysAspLeuSer 65
DB 1 ATCGGGGTGTCTCACACCTTCCTCGTAGAGCTGCACGGCCGCTTGCTGACCTGCTCC 60
QY 66 SerCysAspLeuAlaTrpPheGluGlyArgCysTyrLeuValSerCysProHisLys 85
DB 61 AGCTGTGACCTGGCTGTGGTTCGAGGGCGCTGCTACTGCTGAGCTGCCCCACAAA 120
QY 86 GluAsnCysGluProLysLysMetGlyProIleArgSerTyrLeuThrPheValLeuArg 105
DB 121 GAGAACTGTGAGCCCAAGAAGATGGGCCCATCAGGTCTTATCTCACTTTTGTGCTCCGG 180
QY 106 ProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMetLeuAsnArgGlySer 125
DB 181 CCGTTCAGAGGCTGCACAGCTGCTGGACTATGATGGGACATGATGCTGAACAGGGGCTCC 240
QY 126 ProSerGlyIleTrpGlyAspSerProGluAspIleArgLysAspLeu***PheLeuGly 145
DB 241 CCCTCGGGGATCTGGGGGACTCACCTGAGGATATCAGAAAGACTTGGCCCTTCTAGGC 300
QY 146 LysAspTrpGlyLeuGluGluMetSerGluTyr***AspAspTyrArgGluLeuGluLys 165
DB 301 AAAGATGGGGCTTAGAGGAGATGTCTGAGTACTCAGATGACTACCGGGAGCTGGAGAAG 360
QY 166 AspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAlaGluTyrThrAspTrp 185
DB 361 GAGCTCTTGCAACCCAGTGGCAAGCAGAGAGCCAGAGGGAGTGCCAGGACACGGACTGG 420
QY 186 GlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGlyAspSerProAlaVal 205
DB 421 GGCTACTGCGGGCAGCAGGGGGCTTCACTCCTCTCTGAGACAGCTCCTGGGTG 480
QY 206 ProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAsnGluSerAlaSerThrPro 225
DB 481 CCAGCGGAGACGCAGCAGGACCTGAGCTCCATTACCTGAATGAGTGGCTTCAACCCCT 540
QY 226 AlaProLysLeuProGluArgSerValLeuLeuProLeuProThrThrProSerSerGly 245
DB 541 GCCCAAACTCCCTGAGAGAGTGTGCTTCCCTTGGCCGACTACTCCATCTTCAGGA 600
QY 246 GluValLeuGluLysGluLysAlaSerGlnLeuGlnGlnSerSerAsnSerSerGly 265
DB 601 GAGGTGTGGAGAAAGAAAGGCTTCTCAGCTCCAGGAACAATCCAGCAACAGCTCTGGA 660
QY 266 LysGluValLeuMetProSerHisSerLeuProProAlaSerLeuGluLeuSerSerVal 285
DB 661 AAAGAGGTCTTAATGCCCTCCCATGATCTTCCCTCCCGCAAGCCTGGAGCTCAGTCCAGTC 720
QY 286 ThrValGluLysSerProValLeuThrValThrProGlySerThrGluHisSerIlePro 305
DB 721 ACCGTGGAGAAAGCCAGTGTCTACAGTCAACCCGGGGAGTACAGAGCACAGCATCCCA 780
QY 306 ThrProProThrSerAlaAlaProSerGluSerThrProSerGluLeuProIleSerPro 325
DB 781 ACACCTCCCACTAGCGCAGCCCCCTCTGAGTCCACCCCATCTGAGCTACCCATATCTCCT 840
QY 326 ThrThrAlaProArgThrValLysGluLeuThrValSerAlaGlyAspAsnLeuIle 345

DB 841 ACCACTGCTCCACGAGACA----- 858
QY 346 ThrLeuProAspAsnGluValGluLeuLysAlaPheValAlaProAlaProValGlu 365
DB 859 -----GAA 861
QY 366 ThrThrTyrAsnTyrGluTrpAsnLeuSerHisProThrAspTyrGlnGlyGluIle 385
DB 862 ACAACCTACAACCTATGAATGAATTTAATAAGCCACCCACACAGACTACCAAGGTGAATA 921
QY 386 LysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPhe 405
DB 922 AAACAAGGACACAAAGCAAACTCTTAACCTCTCAATTTGCTCGGACTTTTATGCTTTC 981
QY 406 LysValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLys 425
DB 982 AAAGTCACTGTTTCTAGTGAACACGCTTTGGAGAGGATTTGTCAATGTCACTGCTTAAG 1041
QY 426 ProAlaArgA-gValAsnLeuProProValAlaValValSerProGlnLeuGlnLeu 445
DB 1042 CCT----- 1044
QY 446 ThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThrAspAspThrGluIle 465
DB 1045 -----GAAAGTACAGATGATACATAATA 1068
QY 466 ValSerTyrHisTrpGluGluIleAsnGlyProPheIleGluGluLysThrSerValAsp 485
DB 1069 GTGAGTTATCATTTGGGAAGAAATAAACCGGGCCCTTCATAGAAGAGACATTCAGTTGAC 1128
QY 486 SerProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSerPheArgLeuThrVal 505
DB 1129 TCTCCGCTGTACGCTTGTCTAACCTTGATCCT----- 1161
QY 506 ThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuIleValAsnAsnAlaVal 525
DB 1162 -----GCCCTAATAGTGAACAAATGCTGTG 1185
QY 526 AspTyrProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSer 545
DB 1186 GACTACCCACACGTTGCTTANTGCAGGACCAATACACCAATACCTTTGCCCAAACTCC 1245
QY 546 IleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleValLeuTyrGluTrpSer 565
DB 1246 ATCACTTTGAATGGAAACACAGACGAGTACGATCACCAGATTGCTCTCTATCAGTGGTCC 1305
QY 566 LeuGlyProGlySerGluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeu 585
DB 1306 CTGGGTCTCTGGG----- 1317
QY 586 HisLeuSerAlaMetGlnGluGlyAspTyrThrPheGlnLeuLysValThrAspSerSer 605
DB 1317 ----- 1317
QY 606 ArgGlnGlnSerThrAla***ValThrValIleValGlnProGluAsnAsnArgProPro 625
DB 1318 -----AAAAACAATAGACTCCA 1335
QY 626 ValAlaValAlaGlyProAspLysGluLeuIlePheProValGluSerAlaThrLeuAsp 645
DB 1336 GTGGCTGTGGCCGGCCCTGATAAAGAGCTGATCTTCCCAAGTGAAGTCCACCTCCGAT 1395
QY 646 GlySerSerSerSerAspAspHisGlyIleValPheTyrHisTrpGluHisValArgGly 665
DB 1396 GGGAGCAGCAGCAGCGATGACCCAGCATTTGTTCTACCACTGGGAGCAGCTCAGAGGC 1455
QY 666 ProSerAlaValGluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGln 685
DB 1456 CCCAGTGCAGTGCAGATGGAAATATTGACAAAGCAATAGCCACTGTGTGCTCTCCAG 1515
QY 686 ValGlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGlnLeuSerSerThrSer 705
DB 1516 GTGGGACCTACACTTCCGTTTGACATGTAAGAACCCAGCAGGACTGAGCAGCAGCTCC 1575

QY 706 ThrLeuThrValAlaValLysLysGluAsnAsnSerProProArgAlaArgAlaGlyGly 725
 DB 1576 ACCCTCACTGGCTGTGAAGAAGAAAATAATAGTCTCCAGAGCCCGGCTGGTGGC 1635
 QY 726 ArgHisValLeuValLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAsp 745
 DB 1636 AGACATGTTCTTGCTCCCAATATTCATTTGGATGGTTCAAGGTCTACTGAT 1695
 QY 746 AspGlnArgIleValSerTyrIleTrpIleArgAspGlyGlnSerProAlaAlaGlyAsp 765
 DB 1696 GACCAAGAAATGTTGCTCTATCTGTGGATCCGGATGCCAGAGTCCAGCAGCTGGAGAT 1755
 QY 766 ValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyVal 785
 DB 1756 GTCATCGATGGCTCTGACACAGTGTGGCTCTGACAGCTTACGAATCTGTGGAGGGGGTG 1815
 QY 786 TyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThr 805
 DB 1816 TACATTTTCCACTTCCGAGTACCACAGTCCAGGGGGCTCGGACACAGACACTGCCACT 1875
 QY 806 ValGluValGlnProAspProArgLysSerGlyLeuValGluLeuThrLeuGlnValGly 825
 DB 1876 GTGGAAGTGCAGCCAGACCTTAGGAAGAGTGGCTGTGGAGCTGACCTCGAGGTGGT 1935
 QY 826 ValGlyGlnLeuThrGluGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuLeu 845
 DB 1936 GTTGGCAGCTGCACAGACAGCGGAAGACACCTTGTGTGAGCGAGCTGGCTGTGTGCTG 1995
 QY 846 AsnValLeuAspSerAspIleLysValGlnLysIleArgAlaHisSerAspLeuSerThr 865
 DB 1996 AACGTGCTGGACTCGCATTAAGTCCAGAGATTCGGGGCCACTCGATCTCAGACCC 2055
 QY 866 ValIleValPheTyrValGlnSerArgProPheLysValLeuLysAlaAlaGluVal 885
 DB 2056 GTGATGTTGTTTATGTACAGACAGCGCGCTTTCAAGGTTCTCAAGCTGTGAAGTG 2115
 QY 886 AlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAspPheLeuPheLysVal 905
 DB 2116 GCCCAAAATCTGCATCGGCTCTCAAGAGAGAGCGCGACTTCTTCTGCTTTTCAAGGTC 2175
 QY 906 LeuArgValAspThrAlaGlyCysLeuLysCysSerGlyHisGlyHisCysAspPro 925
 DB 2176 TTGAGGGTGTATACAGAGGTGCTCTGAAGTGTCTGGCCATGGTCACTGGCAGCCC 2235
 QY 926 LeuThrLysArgCysIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyrIle 945
 DB 2236 CTCACAAAGCGCTGATTTGCTCTCACTTATGGATGGAGAACCTTATACAGGTTATATC 2295
 QY 946 TrpAspGlyGluSerAsnCysGluTrpSerIlePheTyrValThrValLeuAlaPheThr 965
 DB 2296 TGGGATGGAGAGCAACTGT----- 2316
 QY 966 LeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCysCysLysArgGlnLys 985
 DB 2316 ----- 2316
 QY 986 ArgThrLysIleArgLysLysThrLysTyrThrIleLeuAspAsnMetAspGluGlnGlu 1005
 DB 2316 ----- 2316
 QY 1006 ArgMetGluLeuArgProLysTyrGlyIleLysHisArgSerThrGluHisAsnSerSer 1025
 DB 2317 -----GGTATCAAGCACCGAAGCACAGACACCACTCCAGC 2352
 QY 1026 LeuMetValSerGluSerGluPheAspSerAspGlnAspThrIlePheSerArgGlnLys 1045
 DB 2353 CTGATGGTATCCGAGTCTGAGTTTGCAGTGTGACAGGACACAACTTTCAGCCGAGAAAG 2412
 QY 1046 MetGluArgGlyAsnProLys 1052
 DB 2413 ATGGAGAGAGGGAATCCAAG 2433

RESULT 4
 BC028869
 LOCUS
 DEFINITION
 Mus musculus expressed sequence AU040320, mRNA (cdna clone
 MGC:28660 IMAGE:4235813), complete cds.
 ACCESSION
 BC028869
 VERSION
 BC028869.1
 KEYWORDS
 GI:20809370
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 4379)
 REFERENCE
 1 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haiech, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shvachenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, V.S., Krzywinski, M.I., Skalka, U., Smallos, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 4379)
 STRAUSBERG, R.
 Direct Submission
 Submitted (01-MAY-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLN)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILLN at: http://image.llnl.gov
 Series: IRAK Plate: 37 Row: 1 Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 23956263.
 Location/Qualifiers
 1. 4379
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:28660 IMAGE:4235813"
 /tissue type="Kidney, normal. 5 month old male mouse."
 /clone_id="NCI CGAP_Kid14"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"

Qy 507 pSerAspGlyAlaThrAsnSerThrAlaAlaLeuLeuValaAsnAlaValAspTy 527
 Db 1664 CTCTGACGGGCTACCAACTCCACCACTGCAAGCTGACTGTGAACAAGCTGTGGACTA 1723
 Qy 527 rProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuLeuProGlnAsnSerIleTh 547
 Db 1724 CCCTCCCGTGGCCAAATGAGAGCCCAACCAAGTAGTACCCCTGCTCAGAACTCATATC 1783
 Qy 547 rLeuAsnGlyAsnGlnSerSerAspHisGlnIleValLeuTyThrGluTrpSerLeuG1 567
 Db 1784 CCTCTTTGAAACACAGAGCAGCGATGACCCAGCATCACAGCTATGATGCTGCCTCAG 1843
 Qy 567 yProGlySerGluGlyHisValValMetGlnGlyValGlnThrProTyThrLeuHisLe 587
 Db 1844 CCCAGCAGCAAGAGGAGGTGGGAGATGACGGAGTTAGAACGCCAGCCCTGCAGCT 1903
 Qy 587 uSerAlaMetGlnGluGlyAspTyThrPheGlnLeuLysValThrAspSerSerArgG1 607
 Db 1904 GTCCGCATGCAAGAGGAGACTATACCTACAGCTACAGTGAAGTACAGCCGAGGACA 1963
 Qy 607 nGlnSerThrAla***ValThrValIleValGlnProGluAsnAsnArgProProValAl 627
 Db 1964 ACAGGCCACGCCCAAGTGAAGTGTGTCAGCCTGAGAACCAACAGCCTCCTCAGGC 2023
 Qy 627 aValAlaGlyProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGlySe 647
 Db 2024 AGATGCAAGGCCACAGCAAGAGCTGACCTCCCGTGGACAGCAACCCCTGGACGGCAG 2083
 Qy 647 rSerSerSerAspHisGlyIleValPheTyThrIleTrpGluHisValArgGlyProSe 667
 Db 2084 CAAGAGCACATGACAGAGAGTCTCTTACTTTGGGAGCAGAGTCCGGGACCTGA 2143
 Qy 667 rAlaValGluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnValG1 687
 Db 2144 CGGGGTGACGAGTGCAGAGTGCACAGCAGGTGCGCCACTGTGACTGGCTGCAAGTCG 2203
 Qy 687 yThrTyHisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrLe 707
 Db 2204 GACTTATGTATTCACCTTGAATGCTCAAGAGATGAGAGAACTACAGAGCCAGAGCTCCGT 2263
 Qy 707 uThrValAlaValLysLysGluLeuAsnSerProProArgAlaAlaGlyAlaGlyArgHi 727
 Db 2264 TAATGTATTGTCAAAGAAAGAAATAAACAACCCAGTACCAAGATCGCTGGGAAGCT 2323
 Qy 727 sValLeuValLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAspG1 747
 Db 2324 GGTGTGACCTTGCCACGACACAGCTGAGCTGGATGGCTCGAGTCTCTCAGATGACAA 2383
 Qy 747 nArgIleValSerTyThrLeuTrpIleArgAspGlyGlnSerProAlaAlaGlyAspValI1 767
 Db 2384 GGGGATAGTCAGCTACCTGTGACTCGAGATGAGACGAGCCAGCCGAGGGAGGTGCT 2443
 Qy 767 eAspGlySerAspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyValTyThr 787
 Db 2444 GAATCACTCTGACCAACCACTGCTCTCTCTCCCACTGTGGAGGGAGCCTACAC 2503
 Qy 787 rPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThrValG1 807
 Db 2504 GTTTCACCTGAAGTGACAGATGCAAGAGGGCAGAGCAGACACCCGAGCAGAGTGGGA 2563
 Qy 807 uValGlnProAspProArgLysSerGlyLeuValGluLeuThrLeuGlnValGlyValG1 827
 Db 2564 AGTGAAGCCTGACCCAGCAAGCAACTAGTGGAGATCATCTTGGATGTGAACGTGAG 2623
 Qy 827 yGlnLeuThrGluGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnVa 847
 Db 2624 TCAGCTGACTGAGAGGCTGAAGGGGAGTCTCATCCGCCAGATCTCTCCCTGGGGGT 2683
 Qy 847 lLeuAspSerAspIleLysValGlnLysIleArgAlaHisSerAspLeuSerThrValI1 867
 Db 2684 GCTGGATCCGACATCATTTGTGCAAAAGATTTCAGCCGTACAGGAGCAGCAACAGAT 2743

Qy 867 eValPheTyThrValGlnSerArgProProPheLysValLeuLysAlaAlaGluValAlaAr 887
 Db 2744 GTTCTTTTGTTCAGAACGACCCCTCCACAGCTCTTCAAGGCCATGAGGTGGCAGC 2803
 Qy 887 gAsnLeuHisMetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuAr 907
 Db 2804 CATGCTCAAGAGCAGCTGCAGAGCAGAGAGCTTCTCTCATCTTCAGAGCCCTGGA 2863
 Qy 907 gValAspThrAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuTh 927
 Db 2864 AATCAGCACAGTACATGTCTGAGTCTGAGTCTGAGCAGTCTGAGTCTGAGTCTGAGTCT 2923
 Qy 927 rLysArgCysIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyThrIleTrpAs 947
 Db 2924 CAACGCTGTGTCTGTGACCCGTTTGGATGGAATTTTCATCAAGGTGTCAGCTGAGGA 2983
 Qy 947 pGlyLysSerAsnCysGluTrpSerIlePheTyThrValThrValLeuAlaPheThrLeuI1 967
 Db 2984 TGGAGACAGCAACTGTGAATGGAGCGTCTCTAGTCTCATCTTCTTGTCTTGTCTTGT 3043
 Qy 967 eValLeuThrGlyGlyPheThrTrpLeuCysIleCysCysLysArgGlnLysArgTh 987
 Db 3044 TGTTCCTTGGGATCCTGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 3100
 Qy 987 rLysIleArgLysLysThrLysTyThrIleLeuAspAsnMetAspGluGlnGluArgMe 1007
 Db 3101 AAAACCCCAAGAGGAAGACAGATACAGATCTCTGATGTCAGAGGCAAAA---GG 3160
 Qy 1007 tGluLeuArgPro-----LysTyGlyIleLysHisArgSerThrGluHisAsnSerSe 1025
 Db 3158 GGAGCTGAACCAACCTCCGAGCAGCAGCAACAGAGAGCCCGCTGAGCAGGAGCAG 3217
 Qy 1025 rLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIlePheSerArgGluLy 1045
 Db 3218 CCTGATGCTTCTGAATCGGAGCTGAGCAGGAC---GATGCCATCTTTCATCATGSCCAGA 3274
 Qy 1045 sMetGluArgGlyAsnProLysValSerMetAsnGlySerIleArgAsnGly 1062
 Db 3275 CCGGAGAGAGGGCAAACTACTGTATGTCAGATGGCTCTGTGCCAAACGGG 3326
 RESULT 5
 LOCUS BC014530 4183 bp mRNA linear PRI 29-JUN-2004
 DEFINITION Homo sapiens polycystic kidney disease 1-like, transcript variant 1, mRNA (cdna clone MGC:11173 IMAGE:3843301), complete cds.
 ACCESSION BC014530
 VERSION BC014530.1 GI:45708392
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4183)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smaluk, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4183)
Strausberg R.
Direct Submission
Submitted (24-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabps-r@mail.nih.gov
Tissue Procurement: AFCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbioology.org>
contact: amadane@systemsbiology.org
Anup Madan, Jessica Fahey, Brin Heiton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 14 Row: m Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1. .4183
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:11173 IMAGE:3843301"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. .4183
/gene="PKD1-like"
/note="synonyms: KIAA1837, PP791"
/db_xref="LocusID:79932"
191..3340
/gene="PKD1-like"
/codon_start=1
/product="polycystic kidney disease 1-like, isoform a"
/protein_id="AAH14530.1"
/db_xref="GI:45708393"
/db_xref="LocusID:79932"
/translation="MEKRLGVKNPNSAWILSGYVWTSAKWLRLSLFLYFTCFCSVLW
LSTDAHSRCQCKTQGVGLRSGENHLLRGTSLSQCLAACQDSACHVFWMLE
GMCIQDCSRQSCAPRTHSSNSMLVFLKFTQADDLGLPDDDPVHLILGLGNWAS
WROSPRAALRAVSSDQSLKIKQKRGSPDVPVPTVTHSKVNDNSNELGLTTS
GSAEWHKATITSSPLTDTLTAELSGGPNVSVQPEISGLATTLPDSTLTLDGSKSS
VPOFAPVASYATPTQAGTSQSTAPYPIKELWSAGESVQITLPKNEVLNAYLVL
EPKGETYTDWOLITHPRDYSQMEGKHSQILKLSKLTPLGYEFKVIYVEGQAHGEP
YNAVTVKPEKRNKPTAIIVSPQOEISLPTSTVLDGSDTDDDKIVYHMEELGKP
LREEKLSDETAILKLSKLVPGNVTFSLTVDSDGATNSTANTLVKAVDPPVANAG
PNQVITLQNSITILFQNSQTDHGIITSYENSLSPSSKGVKVEVGQVPTPLQLSAQGE
GDYTVQITVDTIQGQATQVTVVQENKPPQADAGPKDLTLPDSTLTLDGSKSS
DDQKITSLWEKTVQKQGLQENANSVATVTLGQVGYFTVTLTKDERNLQSSVN
VIVKEENKPEIAKITGNVITLPTSTAEILDGSKSDGKIVSYLWTRDRGSPAAGV
LNHSDHHPILPLNSLVBGTVTFHLTKDAGESDTRTTEVVEKPPDKNNLVLID11
NVSQLTEFLKMGFROIQGLLVGLDSD11VQKIQPYTEOSTKMVFQNEPPEHQIFKG
HEVAMULKSELRKQKADFLIFRALEVNVTVCQLNCSHDGCHDSFTKRCICDPFMWENF
IKVQLRDGDNSCNWSVLYIATFVIVVALGILSWTIVCCCKRQKPKPKRSKYLKLD
ATQDSELELQPTSRAGTIKQKGLLLSSLSMHSSELSDDDAIFTVPDREKGLLHQGQ
SVFNGQTPLKARSPREIL"

REMARK
COMMENT

Alignment Scores:
Pred. No.: 2,02e-120 Length: 4183
Score: 2405.00 Matches: 522
Percent Similarity: 60.13% Conservative: 134

ORIGIN

Qy	314	rGluSerThrProSerGluLeuProIleSerProThrAlaProArgThrValLysGI	334
Db	1078	CCAGAGCACCTCAGACCATACCAGTT- ::: ::: ::: ::: ::: -----ATAAAGGA	1113
Qy	334	uLeuThrValSerAlaGlyAspAenLeuIleThrLeuProAspAenGluValGluLe	354
Db	1114	ACTGGTGGTATCTGCTGGAGAGGTGCCAGATHAACCTGCCTAAAGAATGAAHTTCATT ::: ::: ::: ::: ::: -----	1173
Qy	354	uLysAlaPheValAlaProLaProProValGluThrThrTyraentYrGluTrpAenLe	374
Db	1174	AAATGTCATATGTTCTCCAAGAACCACTTAAGGAGAAAACCTTACACCTACGACTGGCAGCT ::: ::: ::: ::: ::: -----	1233
Qy	374	uIleSerHisProThrAspTyrGlnGlyGluLeuLysGlnGlyHisLysGlnThrLysAs	394
Db	1234	GATTACTCATCTTAGAGACATACGTGGAGAAATGGAAGGGGAAAAATTCCCGACATGCTCAA : ::: ::: ::: ::: -----	1293
Qy	394	nLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAenAl	414
Db	1294	ACTATCGAAGCTCACTCCAGCCCTGTATGAATTCARAAGTGATTGTAGAGGGTCAAAATGC ::: ::: ::: ::: ::: -----	1353
Qy	414	aPheGlyGluGlyPheValAenValThrValLysProAlaAargArgValAenLeuProPr	434
Db	1354	CCATGGGGAAGGCTATGTGAACGTGACAGTCAAGCCAGAGCCCCGTAAAGAATCGGCCCCC ::: ::: ::: ::: ::: -----	1413
Qy	434	oValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIl	454
Db	1414	CATTGCTATTGTGTCACTCAGTTCCAGGAGATCTCTTTGCCAACCACTTCTACAGTCAT ::: ::: ::: ::: ::: -----	1473
Qy	454	eAspGlySerGlnSerThrAspAspThrGluIleValISerTyrHisTrpGluGluIleAs	474
Db	1474	TGATGGCAGTCAAAGCACTGATGATATAAAATCGTTTCAGTACCATTGGGAAGAACTTAA ::: ::: ::: ::: ::: -----	1533
Qy	474	nGlyProPheIleGluGluLysThrSerValAspSerProValLeuargLeuSerAenLe	494
Db	1534	GGGGCCCTCTAAGAGAAGAGAAGATTTCTGAAGATACAGCCATATTAACAATTAAGTAAACT ::: ::: ::: ::: ::: -----	1593
Qy	494	uAspProGlyAsnTyrSerPheArgLeuThrValIThrAspSerAspGlyAlaThrAenSe	514
Db	1594	CGTCCCTGGGAACCTPACACTTTCAGCTTGACTGTAGTAGACTCTGATGGAGCTTCAAATC ::: ::: ::: ::: ::: -----	1653
Qy	514	rThrThrAlaAlaLeuIleValAenAenAlaValAspTyrProProValAlaAenAlaGl	534
Db	1654	TACTACTGCAAACTGTACAGTGAAACAAGCTGTGTGATTACCCCTCTGTGCCAACGCAGG ::: ::: ::: ::: ::: -----	1713
Qy	534	yProAsnHisThrIleThrLeuProGlnAenSerIleThrLeuAenGlyAenGlnSerSe	554
Db	1714	CCCCAACCAAGTGTATCACTCCCTGCCCCCAAACTCCATCACCTCTTTGGGAACCAAGACAC ::: ::: ::: ::: ::: -----	1773
Qy	554	rAspAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHi	574
Db	1774	TGATGATCATGGCATCACCGCTATGAGTGTGTCACTCAGCCACAGCAGCAAGGGAAGT ::: ::: ::: ::: ::: -----	1833
Qy	574	sValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAs	594
Db	1834	GGTGGAGATGCAGGGTGTAGAACACCAACCTTACAGCTCTCTCGCATGCAAGAAGGAGA ::: ::: ::: ::: ::: -----	1893
Qy	594	pTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla**ValTh	614
Db	1894	CTACACTTACCAGTCTCAGTGACTGACACAATATAGGACAGAGCCCACTGCTCAAGTGAC ::: ::: ::: ::: ::: -----	1953
Qy	614	rValIleValGlnProGluAenAenArgProProValAlaValAlaGlyProAspLysGI	634
Db	1954	TGTTATTGTGCNACTGNAACANTATAGCTTCTCAGCGCATGCGGCCCGCATTAAGA ::: ::: ::: ::: ::: -----	2013
Qy	634	uLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSerAspAspHisGI	654
Db	2014	GCTGACCCCTTCTGTGGATAGCAACACCTGGATGGCAGCAAGAGCTCAGATGATCAGAA ::: ::: ::: ::: ::: -----	2073
Qy	654	yIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValGluMetGluAenIl	674
Db	2074	AATTATCTCATCTCTGGGAAAAAACAACAGGGACCTGATGGGTGTCAGCTCGAGATGC ::: ::: ::: ::: ::: -----	2133

Qy	674	eApLyAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyRHiSpheArgLeuTh	694
Db	2134	TAACAGCAGGTGGTACTGTGTGACTGGGTGCAAGTGGGACCTATGTGTTCACCTTGAC	2193
Qy	694	rValLyAspGlnGlnGlyLeuSerSerThrLeuThrValAlaVallyLysG1	714
Db	2194	TGTCAAAGATGAGAGNACTGCAAGCCAGAGCTCTGTGAATGTCTATTGTCAAGAAGA	2253
Qy	714	uAenAsnSerProProArgAlaAlaGlyArgHisValLeuValLeuProAsnAs	734
Db	2254	AATAAACAACACCATATAGCCAAGATAACTGGGAATGTGGTGATACCTACCCACGAG	2313
Qy	734	nSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyLeuTr	754
Db	2314	CACAGCAGACTGGATGGCTCTTAAGTCTCAGATGAACAAGGGGAATAGTCAGCTACCTCG	2373
Qy	754	pIleArgSpGlyGlnSerProAlaGlyAspValIleAspGlySerAspHisSerVa	774
Db	2374	GACTCGAGATGAGGGAGCCACAGCAGGGGAGGTGTAAATCACTCTGACCATCACCC	2433
Qy	774	lAlaLeuGlnLeuThrAsnLeuValGluGlyValTyThrPheHisLeuArgValThrAs	794
Db	2434	TATCCTTTTCTTCAACACCTGTGTGAGGAACCTACACTTTTTCACCTGAAGAGTACCGA	2493
Qy	794	pSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArgLy	814
Db	2494	TGCAAAAGGGTGGAGGTGCACACAGCCGGACCACTGTGGAGGTGAAAACCTGATCCACGAA	2553
Qy	814	sSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLy	834
Db	2554	AAACMACTGGTGGAGATCATCTTGGATATCAACGTCACTGACGTCACTGAGAGGCTGAA	2613
Qy	834	sAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysVa	854
Db	2614	GGGGATGTTTCATCGCCAGATTGGGTCTCTCTGGGGTCTGGATTCGACATCATTTGT	2673
Qy	854	lGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTyRValGlnSerAr	874
Db	2674	GCAAAAGATTTCAGCCGTACACGGAGCAGACCAAAATGGTATTTTTTTGTTCAAAACGA	2733
Qy	874	gProPropheLysValLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSe	894
Db	2734	GCCTCCCAACAGATCTTCAAGGCCATGAGGTGGCAGCGATGCTCAAGAGTGAGCTGGC	2793
Qy	894	rLysGlyLysAlaAspPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLe	914
Db	2794	GAAGCAAAAGGCAGACTTTTGTATTCAGAGCCTTGGAGTCAACACTGTCAATGTCA	2853
Qy	914	uLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCysIleCysSerHi	934
Db	2854	GCTGAACGTGTTCCGACCATGGCCACTGTGACTCGTTCCCAACCGCTGTATCTGTGACCC	2913
Qy	934	sLeuTrpMetGluAsnLeuIleGlnArgTyRlIeTrpAspGlyGluSerAsnCysGluTr	954
Db	2914	TTTTTGTATGGAGAAATTTCATGAAGTGCAGCTGAGGGATGGAGACGACCACTGTGAGTG	2973
Qy	954	pSerIlePheTyRValThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheTh	974
Db	2974	GAGCGTGTATATGTTATCATCTGCTACCTTTGTCTATTGTGTTCCTTGGGAATCCTGTC	3033
Qy	974	rTrpLeuCysIleCysCysLysArgGlnLysArgThrLysIleArgLysLysThrLy	994
Db	3034	TTGACTGTGATCTGTTGTTGTAAGAGGCAAAA---GGAAAAACCCCAAGAGAAAAAGCAA	3090
Qy	994	sTyThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro-----Ly	1012
Db	3091	GTACAAGATCTGTGATGCCCGGAT---CAGGAAGCCCTGAGCTGAAGCAACCTCCCG	3147
Qy	1012	sTyRlyLyleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGluSerGl	1032
Db	3148	AGCAGGCGATCAACAGAAAGAGCCCTTTTGTCTAAGTAGCAGCCTGATGCATCCGAGTCAGA	3207
Qy	1032	uPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnProLy	1052

Db 3208 GCTGACAGCGAT---GATGCCATCTTTACATGGCCAGACCAGAGAGAGGCAAACTCCT 3264
 Qy 1052 eValSerMetAsnGlySerIleArgAsnGly 1062
 Db 3265 GCATGGTCAGAAATGGCTCTGTACCAACGGG 3295

RESULT 6

AX195570 3908 bp DNA linear PAT 28-AUG-2001
 LOCUS
 DEFINITION Sequence 15 from Patent WO0151636.
 ACCESSION AX195570
 VERSION AX195570.1 GI:15386083

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

1 Lal, P., Bandman, O., Patterson, C., Walla, N.K., Nguyen, D.B., Yue, H.,

Khan, F.A., Tang, Y.T., Baughn, M.R., Lu, D.A., Yang, J., Burford, N.,

Au-Young, J. and Reddy, R.

Secreted proteins

Patient: WO 0151636-A 15 19-JUL-2001;

Incyte Genomics, Inc. (US)

Location/Qualifiers

1..3908

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="Incyte ID No: 2323763CB1"

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 2.1e-120 Length: 3908
 Score: 2404.00 Matches: 522
 Percent Similarity: 60.13% Conservative: 134
 Best Local Similarity: 47.85% Mismatches: 318
 Query Match: 43.08% Indels: 117
 DB: 6 Gaps: 17

US-10-764-390-3 (1-1072) x AX195570 (1-3908)

Qy 4 ProThrGlyValLeuSerSerLeuLeuLeu-----LeuValThrIleAlaGlyCys-Al 21
 Db 296 CTGTACCTGTTTATCTACTGCTTTTGTCTACGGTCTGTGTTGTCAACAGATGCCAG 355
 Qy 21 aArgLysGlnCysSerGluGlyArgThrTyrSerAsnAlaValIleSerProAsnLeuGl 41
 Db 356 TGAGAGCAGGTGCCAGCAGGGGAGACACAAATTTGGAGTTGGCTGAGATCTGGGGGAGA 415
 Qy 41 uThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAlaAlaCy 61
 Db 416 A---AATCACTCTGCTGCTTTCTTGAAGGAACCCCTCTCTCCAGTCATGTGGGCTGCCTG 472
 Qy 61 sCysAspLeuSerSerCysAspLeuAlaTrpPheGluGlyArgCysTyrLeuValSe 81
 Db 473 CTGCCAGGACTCTGCTCGCATGTCCTTTTGGTGCTAGAGGGATGTGCATTCAGGCAGA 532
 Qy 81 rCysProHisLysGluAsnCysGluProLysLysMetGlyProIleArgSerTyrLeuTh 101
 Db 533 CTGCAGCAGGCCCCAGAGCTGCCGGCTTTTAGGCACACACTCTCCAAATTCATGCTGCT 592
 Qy 101 rPheValLeuArgProValGlnArgProAlaGlnLeuLeuLeuAspTyrGlyAspMetMet 121
 Db 593 GTTT---TTAAAAAAATTCCAAACTGCAGATGATTG----- 626
 Qy 121 uAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAspIleArgLysAspLe 141
 Db 627 -----GCCTTTCTACTCTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 658
 Qy 141 u***PheLeuGlyLysAspTrpGlyLeuGluMetSerGluTyr***AspAspTyrAr 161

Db 659 TCTGGGCTAGTGTGAACCTGGCA----- 683
 Qy 161 gGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAlaGl 181
 Db 684 -----TCTTGGAGGCGAGCCCAACCCAGAGCTGCCT 715
 Qy 181 uTyrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGlyAs 201
 Db 716 C----- 716
 Qy 201 pSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAsnGluSe 221
 Db 717 -AGACCTGCTGTCTCTCCAGTCAGCAGCAG----- 746
 Qy 221 rAlaSerThrProAlaProLysLeuProGluArg-----SerValLeuLeuPr 237
 Db 747 -----AGCTTAATCAGGAAGCTTCAGAGAGAGAGGTAGTCCAGTCAGTGTATACACC 799
 Qy 237 oLeuProThrThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGl 257
 Db 800 TATAGTGACACAGCATTTCT-----AAAGTGAATGACTCCAACGAATTAGGTGGTCTGAC 853
 Qy 257 nGluGlnSerSerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProPr 277
 Db 854 TACCAGTGGCTCTGCAGAGGTCCACAGGCGATTACAATTTCCAGTCCCTTAACACACAGA 913
 Qy 277 oAlaSerLeuGluLeuSer-----SerValThrValGluLysSerPr 291
 Db 914 CTGACTGCAGAGCTGTCTGGTGGGCCAAAGAAATGATCAGTGCACCTGAAATATACAGA 973
 Qy 291 oValLeuThrValThrProGlySer-----ThrGluHisSe 303
 Db 974 GGGTCTTGCTACTAGCCCCAGCACTCAACAAGTAAAGATTCTGAGAAAAACCCAGATTGC 1033
 Qy 303 rIleProThrPro-----ProThrSerAlaAlaProSe 314
 Db 1034 TGTCCCCAGCCAGTGGCTCCCTCCTACAGTTATGCTACCTACCCCTCCCGAGGCTCTTT 1093
 Qy 314 rGluSerThrProSerGluLeuProIleSerProThrThrAlaProArgThrValLysGl 334
 Db 1094 CCAGAGCACCTCAGCACCATACCCAGTT-----ATTAAGA 1129
 Qy 334 uLeuThrValSerAlaGlyAspAsnLeuIleIleThrLeuProAspAsnGluValGluLe 354
 Db 1130 ACTGCTGTATCTGCTGGAGAGGTCCAGATAACCTGCCTAAGATGAAGTTCAATT 1189
 Qy 354 uLysAlaPheValAlaProAlaProProValGluThrThrTyrAsnTyrGluTrpAsnLe 374
 Db 1190 AAATGCATATGTTCTCCAAGAACCACTTAAAGGAGAAACCTACACCTACGACTGGCAGCT 1249
 Qy 374 uIleSerHisProThrAspTyrGlnGlyLysGlnGlyHisLysGlnThrLeuAs 394
 Db 1250 GATTACTCATCTCAGAGACTAGTGGAGAAATGGAAGGAAACATTCACGATCTCAA 1309
 Qy 394 nLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAsnAl 414
 Db 1310 ACTATCGAAGCTCACTCCAGGCTCTGATGAATTCAGAGGTGTAGAGGGTCMAAATGC 1369
 Qy 414 aPheGluGluGlyPheValAsnValThrValLysProAlaArgArgValAsnLeuProPr 434
 Db 1370 CCATGGGAAAGGCTATGTGAACGTGACAGTCAAGCCAGAGCCCGTGAAGATCGGCCCCC 1429
 Qy 434 oValAlaValValSerProGlnLeuGlnLeuThrLeuProLeuThrSerAlaLeu1 454
 Db 1430 CATTCCTATGTGTCACTTCAGTCCAGGAGATCTTTTGGCCCAACCACTTCTACAGTCAT 1489
 Qy 454 eAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHisTrpGluGluLeAs 474
 Db 1490 TGATGGCAGTCAAAGCACTGATGATGATAAATCGTTAGTACCTACCTTGGGAAGAACTAA 1549
 Qy 474 nGlyProPheIleGluGluLysThrSerValAspSerProValLeuArgLeuSerAsnLe 494
 Db 1550 GGGGCTCTTAAGAGAGAAGAGAGATTCTTGAAGATACAGCCATATTAAATAAGTAAACT 1609

Qy	494	uAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsn	514
Db	1610	CGTCCCTGGCAACTACACTTTACGTTTGACTGTAGTAGACTCTGTATGGAGCTACCAACTC	1669
Qy	514	rThrThrAlaAlaLeuLeuValAenAenAlaValAspTyrProProValAlaAenAlaGl	534
Db	1670	TACTACTGCAAAACCTGCAGGTGAACAAGCTGTGGATTACCCCTCTGGCCAAACGCGAG	1729
Qy	534	yProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSer	554
Db	1730	CCCCAACCAAGTGATCACTTGGCCCCAAATACTCATCATCCCTCTTTGGAAACCAAGCAC	1789
Qy	554	rAspAspHisGlnIleValLeuTyrGluTyrSerLeuGlyProGlySerGluGlyIysH	574
Db	1790	TGATGATCATGGCATCACCAAGCTATGAGTGGTCACTCAGCCCAAGCAGCAAGGGAAGT	1849
Qy	574	sValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAs	594
Db	1850	GGTGGAGATCAGGGGTGTAGAACACCAACCTTACAGCTCTCTGCGATGCAAGAAGAGA	1909
Qy	594	pTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla**ValTh	614
Db	1910	CTACACTTACGAGCTCAGTGAAGTGAACAATAGGACAGAGGCCACTGCTCAAGTGAC	1969
Qy	614	rValIleValGlnProGluAenAenArgProProValAlaValAlaGlyProAspLysGl	634
Db	1970	TGTTATTGTGCAACTGAAAAACAATAAGCTCTCTCAGCGCAGATGCAGGCCACGATAAGA	2029
Qy	634	uLeuIlePheProValGlnSerAlaThrLeuAspGlySerSerSerSerAspAspHisGl	654
Db	2030	GCTGACCTCTCTGTGATGACCAACCTCGATGGCAGCAGCAGCTCAGATGATCAGAA	2089
Qy	654	yIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValGluMetGluAenIl	674
Db	2090	AAITATCTCATATCTCTGGGAAAAACACAGGGACCTGATGGGGTGCGAGCTCGAAGATGC	2149
Qy	674	eAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuTh	694
Db	2150	TAACAGCAGTGTGCTACTGTGACTGGCTGCTCAAGTGGGACCTATGTGTTCACCTTGAC	2209
Qy	694	rValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAlaValLysLysGl	714
Db	2210	TGTCAAGATGAGAGGAACCTGCAAGGCCAGAGCTCTGTGAATGTCATTGTCAAGAAGA	2269
Qy	714	uAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAs	734
Db	2270	AATAAACAAACCACTATAGCAAGATAACTGGGAATGTGTGTGATACCTTACCCTACCGAG	2329
Qy	734	nSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrLeuTr	754
Db	2330	CACAGCAGACTGTGATGGCTTAAGTCTCAGATGACARGGAATAGTCACTACCTCTG	2389
Qy	754	pIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerVa	774
Db	2390	GACTCGAGATGAGGGGAGCCCGACGAGCAGGGGAGGTGTTAAATCACTCTGACCATCACCC	2449
Qy	774	lAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThrAs	794
Db	2450	TATCCTTTTCTTTCAAACTGGTGTGAGGGAACTTACACTTTTCCACTGAAAGTGACCGA	2509
Qy	794	pSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProAsGly	814
Db	2510	TGCAAGGGGTGAGAGTGACACAGACCCGACCACTGTGGAGGTGAACCTGATCCAGGAA	2569
Qy	814	sSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLy	834
Db	2570	AAACAACTGTGTGAGATCATCTTTGGATATCAACGTCAGTCAGCTAACTGAGGCTGAA	2629
Qy	834	sAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysVa	854
Db	2630	GGGGATGTTTCATCCGCCAGATTGGGGTCTCTCTGGGGGTGCTGGATTCCGACATCATGT	2689

Qy	854	lGlnLysIleA	ArgAlaHisSerAspLeuSerThrValIleValPheTyrValGlnSerAr	874
Db	2690	GCAAAAGATT	CAGCCGTCACAGGAGCAGACCAAAAATGGTATTTTGTGTTCAAAACGA	2749
Qy	874	gProPheLysValLeuLysAla	AGluValAlaArgAsnLeuHiSwetArgLeuSe	894
Db	2750	GCCTCCCCAC	CAGATCTTCAAGGCCATGAGGTGGCAGCGATCTCAAGAGTCAGTGGC	2809
Qy	894	rLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLe	914	
Db	2810	GAAGCAAAAGGCAGAC	CTTTTGTATATTACAGAGCCTTGGAGTCAACACTGTACATGTCA	2869
Qy	914	uLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCysIleCysSerHi	934	
Db	2870	GCTGAAC	TGTTCGACCATGCCACTGTGACTCGTTTCAACAAACGCTGTATCTGTGACCC	2929
Qy	934	sLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyGluSerAsnCysGluTr	954	
Db	2930	TTTTTGGATGGAGAA	TTTCAATCAAGGTGCAGCTGAGGATGGAGACAGCAACTGTGATG	2989
Qy	954	pSerIlePheTyrValThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheTh	974	
Db	2990	GAGCGTGTATAT	TATCATCTGCTACCTTTGTCATTGTTGTTGCTTGGGAATCCTGTC	3049
Qy	974	rTrpLeuCysIleCysCysCysLysArgGlnLysArgThrLysIleArgLysLysThrLy	994	
Db	3050	TTTGACTGTGATCTGTTGTTGTTGAAGAGGCAAAA	--GGAAAAACCAAGAGGAAAAACAA	3106
Qy	994	sTyrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro	1012	
Db	3107	GTACAAGATCTGTGATGCCAGGAT	---CAGGAAGCCTGGAGCTGAAGCAACCTCCCG	3163
Qy	1012	sTyrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGluSerGl	1032	
Db	3164	AGCAGGATCAACAGAAAGCCCTTTTGTCTAAGTAGCAGCGCTGATGCATCCGAGTCAG	3223	
Qy	1032	uPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnProLy	1052	
Db	3224	GCTGCACAGCGAT	---GATGCCATCTTTACATGCCACCGAGAGAGGCAAACTCCT	3280
Qy	1052	sValSerMetAsnGlySerIleArgAsnGly	1062	
Db	3281	GCATGGTCAGAAATGGCTCTGTACCCACGGG	3311	
RESULT 7				
LOCUS	AV163234		4000 bp mRNA linear	PRI 04-NOV-2002
DEFINITION			Human sapiens polycystic kidney disease 1-related protein mRNA,	
			complete cds.	
ACCESSION	AV163234			
VERSION	AV163234.1		GI:24559835	
KEYWORDS				
SOURCE			Human sapiens (human)	
ORGANISM			Human sapiens	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.	
TITLE			1 (bases 1 to 4000)	
JOURNAL			Hao, D. and Hooi, S.	
FEATURES			Direct Submission	
			Submitted (12-OCT-2002) Physiology, National University of	
			Singapore, 2 Medical Drive, Singapore 117597, Singapore	
			Location/Qualifiers	
			1..4000	
			/organism="Homo sapiens"	
			/mol_type="mRNA"	
			/db_xref="taxon:9606"	
			/chromosome="1"	
			/map="1p31"	
			208..3357	
			/note="PKD1-like protein; membrane protein"	
			/codon_start=1	
			/product="polycystic kidney disease 1-related protein"	
			/protein_id="AA061054.1"	
CDS				

/db_xref="GI:24559836"
/translation="MEKRLGVKPNPASPILSGYVWQTSKAWLSRLSLFYVFCPSVLW
LSTDAESRCCOQKTFQGVLSGGENHLMLEGTPLSQSCWAAACQDSACHVFWML
GMCIOADCSRQSCRAFTHSHSNMLVFLKFKQFADDLGLFPEDDVPHLLGLGNWAS
WQSPRAALRPVSPSSDOQLIRKLQRKSPSDVVPIVTOHSHKVNDSNELGGUTTS
GRSAEYKAITTSPLTDLTAELSGPKPNVQPEISRELAATTPSQVSKSEKQTIA
VPOVPASYSYATPQASFOSTAPYVKELVVVSAGESQITLFPKNEVLQNAVYLO
EPKGETVYDWOLITHPRDYSGEMGHKSILKSLTGPLKEPKVVEGONAHGEV
YVNVTKPEPRKRPPIAIVSPOFOEISLPTTSTVLDGQSOTDDDKIVOVHWEELKGP
LNEKISDELTALKLSKLVPGNYTSLTVSDSGATSTTANLTNKAIVDPYPPVANAG
PNOVITLSDNTLFGNSDTHGITSYSSPKSGKGVEMQVTRPTIQLSAMQE
GDYTYQLVTRTQGOATAQTVIVPENNPQADAGPDKELTLPVDSITLDGSSK
DDQKLIISWEKTQPGDQVLENANNSVATVGLQVTVYVFTLVKDBRNILDSQSSV
VIVKEINPKPTAKTGNVITLPTSLAELDGSKSSDDKIVSYLWTRDEGSPAAGEV
LNHSDHPHILSLNVECTYFHLKVTADKESDTRTVEVKPDPRKNLVEILLDI
NVSQUTERLKGFIHQIGVLGLVSDIIVKIQYITEQSTQWFFVQNEPPIQFKG
HEVAMLKSELRKQKADLFILFALENVNTVQOLNCSHDHCDSTFKRCICDPFFWNEF
IKVQURDGSNCENSVLXVITATFVIVALGILSWTVICCKRKQKPKRKSXYKILD
ATDOESLELKPTSRAGIKQKGLILLSLMSHSELDSDDAITFWPDREKGLLHGQNG
SVPNQGTPLKARSPREELI"

ORIGIN

Alignment Scores:

Pred. No.: 2,16e-120 Length: 4000
Score: 2404.00 Matches: 522
Percent Similarity: 60.13% Conservative: 134
Best Local Similarity: 47.85% Mismatches: 318
Query Match: 43.08% Indels: 117
DB: 9 Gaps: 17

US-10-764-390-3 (1-1072) x AV163234 (1-4000)

QY 4 ProThrGlyValLeuSerSerLeuLeuLeu-----LeuValThrIleAlaGlyCys-Al 21
DB 297 CCTGTACCTGTTTATCTACTTCTTCTGCTTCAGCGTTCTGTGTTGTCACAGATGCCAG 356
QY 21 aArgLysGlnCysSerGluGlyArgThrTyrSerAsnAlaValIleSerProAsnLeuG1 41
DB 357 TCAGAGCAGGTGCCAGCAGGAGGAGACACAAATTTGGAGTTGGCTTGAGATCTGGGGGAGA 416
QY 41 uThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAlaAlaCy 61
DB 417 A---AATCACCTGTGGCTTCTTGAAGGAACCCCTCTCTCCAGTCATGTTGGGCTGCCTG 473
QY 61 sCysAspLeuSerSerCysAspLeuAlaTrpPheGluGlyArgCysTyrLeuValSe 81
DB 474 CTGCCAGGACTCTGCCCTGCATGTCCTTTGGTGGCTAGAGGATGTGCATTCAGGCAGA 533
QY 81 rCysProHisLysGluAsnCysGluProLysLysMetGlyProIleArgSerTyrLeuTh 101
DB 534 CTGCAGCAGCCCGCAGAGCTGCGGGCTTTTAGGACACACTCTCCAATTCATGCTGTGT 593
QY 101 rPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMetIe 121
DB 594 GTTTT---TTAAAAAATTCCAAACCTCAGATGTTG----- 627
QY 121 uAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAspIleArgLysAspIe 141
DB 628 -----GGCTTTCTACCTGAAGATGATGTACCATCT 659
QY 141 u**PheLeuGlyLysAspTrpGlyLeuGluMetSerGluTyr**AspAspTyrAr 161
DB 660 TCTGGGGCTAGGTGGAACCTGGCA----- 684
QY 161 gGluLeuGluLysAspLeuGlnProSerGlyLysGlnGluProArgGlySerAlaGl 181
DB 685 -----TCTTGGAGGCAGAGCCCAACCCAGAGCTGCAC 716
QY 181 uTyrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGlyAs 201
DB 717 C----- 717
QY 201 pSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAsnGluSe 221
DB 718 -AGACCTGCTGTATCTCTCCAGTGCACAGCAG----- 747
QY 221 rAlaSerThrProAlaProLysLeuProGluArg-----SerValLeuLeuPr 237
DB 748 -----AGCTTAATCAGGAAGCTTCAGAAAGAGAGGTAGTCCCATGACGTAGTAGTACACC 800
QY 237 oLeuProThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuG1 257
DB 801 TATAGTGACACAGCATCT-----AAAGTGAATGACTCCAACGAATTAGGTGGTCTGAC 854
QY 257 nGluLnsSerSerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProPr 277
DB 855 TACCAGTGGCTCTGCAGAGGTCCCAAGGCGATTACAAATTTCCAGTCCCTCCCAACACAGA 914
QY 277 oAlaSerLeuGluLeuSer-----SerValThrValGluLysSerPr 291
DB 915 CTTGACTGCAGAGCTGTCTGGTGCCCAAGAATGTATCAGTGCACCTGAAATATACAGA 974
QY 291 oValLeuThrValThrProGlySer-----ThrGluHisSe 303
DB 975 GGGTCTTGCTACTAGCCCCAGCCTCAACAAGTAAAGTTCTGAGAAAACCCAGATTGC 1034
QY 303 rIleProThrPro-----ProThrSerAlaAlaProSe 314
DB 1035 TGTCCCCCAGCAGTGGCTCCCTCCTACAGTATTATGCTACCCCTACCCCCAGGCTCTTT 1094
QY 314 rGluSerThrProSerGluLeuProIleSerProThrThrAlaProArgThrValLysGl 334
DB 1095 CCAGAGCACCTTCAGCACCATATCCCGAGT-----ATAAAGGA 1130
QY 334 uLeuThrValSerAlaGlyAspAsnLeuIleIleThrLeuProAspAsnGluValGluLe 354
DB 1131 ACTGGTGTATCTGCTGGAGAGAGTCCAGATACCTGCTTAAGATGAATGAATTCAT 1190
QY 354 uLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTyrGluTrpAsnLe 374
DB 1191 AAATGCATATGTTCTCCAGAAACACCACTAAAGGAGAAAACCTACACTACGACTGCAGCT 1250
QY 374 uLleSerHisProThrAspTyrGlnGlyGluLysGlnGlyHisLysGlnThrLeuAs 394
DB 1251 GATTACTCATCTAGAGACTACAGTGGAGAAATGGAAGGAAACAAATCCCAATCCTCAA 1310
QY 394 nLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAsnAl 414
DB 1311 ACTATCGAAGTCACTCCAGGCTGTATGAATTCAAAGTGTGTAGAGGTCAAAATGC 1370
QY 414 aPheGlyGluGlyPheValAsnValThrValLysProAlaArgArgValAsnLeuProPr 434
DB 1371 CCATGGGGAAGGCTATGTGAACGTGACAGTCAAGCCAGAGCCCGCTAAGAATCGGCCCC 1430
QY 434 oValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuI 454
DB 1431 CATTCCTATTGTGTCACTTCCAGGAGATCTCTTTGGCAACCACTTCTACAGTCAT 1490
QY 454 eAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHisTrpGluGluLeAs 474
DB 1491 TGATGGCAGTCAAGCACTGATGATGAATAATCGTTCAATACCATTTGGGAAGACTTAA 1550
QY 474 nGlyProPheLeuGluGlyLysThrSerValAspSerProValLeuArgLeuSerAsnLe 494
DB 1551 GGGGCTCTAAGAGAAGAGAGATTTCTGAAGATACAGCCATATAAATAAGTAAACT 1610
QY 494 uAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSe 514
DB 1611 COTCCCTGGGAACACTACATTTTCAGCTGTAGTAGTCTGTATGGAGTACCAACTC 1670
QY 514 rThrThrAlaAlaLeuLeuValAsnAsnAlaValAspTyrProProValAlaAsnAlaGl 534
DB 1671 TACTACTGCAAACTGCAGTGAGAAACAAAGCTGTGATTTACCCCTCTGTGCCCAACGCG 1730
QY 534 yProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSe 554

```
Db 1731 CCCCAACCAAGTGATCACCTGCCCCAAAACCTCCATCACCCCTCTTTGGGAACCAAGAGCAC 1790
Qy 554 rAspAphisGlnIleValLeuYrGluTrpSerLeuGlyProGlySerGluGlyLysEhI 574
Db 1791 TGATGATCATGGCATCACAGCTATAGTGCTCACTCAGCCCAAGCAGCAAGGAAAGT 1850
Qy 574 sValValMetGlnGlyValGlnThrProTyLeuHiSLeuSerAlaMetGlnGluGlyAs 594
Db 1851 GGTGAGATGAGGGTGTTAGAACCAACCTTACAGCTCTCTCGATGCAAGAGGAGA 1910
Qy 594 pTyTrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla***ValTh 614
Db 1911 CTACACTTACAGCTCACAGTGACTGACACATAGGACAGCAGGCCACTGCTCAAGTGAC 1970
Qy 614 rValIleValGlnProGluAenAenArgProProValAlaValAlaGlyProAspLysG1 634
Db 1971 CGTTATTGTGCAACCTGAAACCAATAAGCCTCCTCAGGACAGATGCGAGCCAGATAAGA 2030
Qy 634 uLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSerAspAspHisG1 654
Db 2031 GCTGACCCCTTCTGTGGATAGCAACCTCGATGGCAGCAGAGACTCAGATGATCAGAA 2090
Qy 654 yIleValPheTyTrHisTrpGluHiSValArgGlyProSerAlaValGluMetGluAenI1 674
Db 2091 AATTATCTCATCTCTGGGAAAAACACAGGACCTGATGGGTGCGAGCTCGAATGC 2150
Qy 674 eAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyHisPheArgLeuTh 694
Db 2151 TAACAGCAGTGTGTACTGTGCTGAGTGGGCTGCAAGTGGGACCTATGTGTTCACCTTGC 2210
Qy 694 rValIleAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAlaValLysG1 714
Db 2211 TGTCAAGATGAGAGAACCTGCAAGCCAGAGCTCTGTGAATGTCTGTTGCAAGAAGA 2270
Qy 714 uAsnAsnSerProProArgAlaArgAlaGlyGlyVargHisValLeuValLeuProAsnAs 734
Db 2271 AATAAACAACCACTATAGCAAGATAACTGGGNAATGTGTGATTACCTTACCACAG 2330
Qy 734 nSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyLeuTr 754
Db 2331 CACAGCAGCTGGATGCTCTAAGTCTCAGATGACAAGGGAATAGTCAGCTACCTCTG 2390
Qy 754 pIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerVa 774
Db 2391 GACTCGAGATGAGGGAGCCCGCAGCAGCAGGGAGGTGTTAAATCACTCTGACCATCACCC 2450
Qy 774 lAlaLeuGlnLeuThrAsnLeuValGluGlyValTyThrPheHisLeuArgValThrAs 794
Db 2451 TATCTCTTTTCTTTCAAACTGGTTGAGGGAACTTACACTTTTTCACCTGAAGTGACCGA 2510
Qy 794 pSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArgLy 814
Db 2511 TGCAAGGGTGAGAGTGACACAGACCGGACCACTGTGGAGGTGAACCTGATCCAGGAA 2570
Qy 814 sSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLy 834
Db 2571 AAACAACCTGTGGAGATCATCTTGGATATCAACGTCACTGAGCTAACTGAGAGCTGAA 2630
Qy 834 sAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysVa 854
Db 2631 GGGGATGTTCCCGCCAGATTTGGGTCTCTCTGGGGTGTCTGGATTCGACATCATTTGT 2690
Qy 854 lGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTyTrValGlnSerAr 874
Db 2691 GCAAAAGATTGAGCGGTACACGGACAGCAGCAGCAACCAAAATGTTATTTTGTTCAAAACGA 2750
Qy 874 gProProPheLysValLeuLysAlaAlaGluValAlaAlaArgLeuHisMetArgLeuSe 894
Db 2751 GCCTCCCAACAGATCTTCAAGGCGCATCAGGTGGCAGCGATGCTCAAGAGTGAGCTCGC 2810
Qy 894 rLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLe 914
Db 2811 GAAGCAAAAGGAGAGACTTTTGTATATTAGAGCCTTGGAAAGTCAACACTGTCACATGTCA 2870
```

```
Qy 914 uLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCysIleCysSerHi 934
Db 2871 GCTGAACCTGTTCCGACCATGGCCACTGTGACTCGTTCCACCAACGCTGTATCTGTGACCC 2930
Qy 934 sLeuTrpMetGluAsnLeuIleGlnArgTyTrIleTrpAspGlyGluSerAsnCysGluTr 954
Db 2931 TTTTGGATGGAGAAATTTTCATCAAGGTGCAGCTGAGGATGGAGACACCACTGTGAGTG 2990
Qy 954 pSerIlePheTyTrValThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheTh 974
Db 2991 GAGCGTGTATATGTTATCATCTTGTCTGCTTGTCTGCTTGTCTGCTTGTCTGCTGCTGCTG 3050
Qy 974 rTrpLeuCysIleCysCysCysLysArgGlnLysArgThrLysIleArgLysLysThrLy 994
Db 3051 TTGGACTGTGATCTGTGTTGTAAAGAGCAAAA---GGAATAACCCAGAGAGAAAGCAA 3107
Qy 994 sTyTrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro-----Ly 1012
Db 3108 GTACAAGATCTGGATGCCCGAT---CAGAAAGCTGGAGCTGAAGCCAACTCCCG 3164
Qy 1012 sTyTrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGluSerG1 1032
Db 3165 AGCAGGCATCAACAGAAAGGCTTTTGTCTAAGTAGCAGCCTGATGCACTCCGAGTCAGA 3224
Qy 1032 uPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnProLy 1052
Db 3225 GCTGCAGCAGCGAC---GATGCCATCTTTACATGCCAGACCCAGAGAGAGGCGAACTCCT 3281
Qy 1052 sValSerMetAsnGlySerIleArgAsnGly 1062
Db 3282 GCATGGTCAGATGGCTCTGTGTACCAACGGG 3312

RESULT 8
AF275679 Homo sapiens PP791 protein mRNA, complete cds.
LOCUS AF275679 4001 bp mRNA linear PRI 15-JAN-2002
DEFINITION AF275679 Homo sapiens PP791 protein mRNA, complete cds.
ACCESSION AF275679
VERSION AF275679.1 GI:10880790
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4001)
AUTHORS Qin,W.X., Wan,D.F., Zhou,X.M., Zhang,P.P., Jiang,H.Q., Huang,Y.,
Zhao,X.T. and Gu,J.R.
Novel human cDNA clones with function of inhibiting cancer cell
growth
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4001)
AUTHORS Qin,W.X., Wan,D.F., Zhou,X.M., Zhang,P.P., Jiang,H.Q., Huang,Y.,
Zhao,X.T. and Gu,J.R.
Direct Submission
JOURNAL Submitted (07-JUN-2000) National Laboratory For Oncogenes & Related
Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Fu Road, Shanghai
200032, P.R. China
FEATURES
Location/Qualifiers
source 1..4001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
208..2130
/codon_start=1
/product="pp791 protein"
/protein_id="AAG24389.1"
/db_xref="GI:10880791"
/translaton="MEKRLGVKPNPASWILSGYTWQTSAKMLRSLYLEPTVFCPSVLW
LSTDASRSQQQKTFQVGLRSGGNHLWLEGTPSLQSCWAAACQDSACHVFWMLE
GMCTQADCSRQPSQCRAPTRSSNMLVFLKKFTQADDLGFLPDDVDVPHLLGLGWNWAS
WROSPRAALRLPAVSSDQSLIRLKRKGSFSDVTPITVQYSKVDNSNLGGLTTS
GSAPVHKATITSSPLTDLTALSLGSKNSVQPEISEGLATPTSTQOVKSSEKQIA
VQPVPAPSYSTATPTPOASPOSTSAPFVVKELVVSAGESVQIILPKNEVQLNAVLYL
```

EPPKGITYYDQLIHTPRDYSGEMEGKHSQILKSLKLTPLGLYEFKVIYEGQNAHGBG
 YVNVTKPEPRKNRPPIAIVSFOEISLPTTSTVIDGSQSDDDDKIYQYHWEILKGP
 LREKISDITAILKSLKVPNGYTFSLTVVDSGDGNTSTTANLTWNKAVDYPVPVANAG
 PNQVITLQNSSTLFGNSTDDHGITSYENSLSRSSKGVEMQGVRRPTLQLSAMQE
 GDYITQIVTDITGQQAQAVTVIIVQPENNKFPQADAGDKELTULFVDSITIDGSKS
 DDQKIISYLWEKNTGT"

ORIGIN

Alignment Scores:
 Pred. No.: 2,29e-119 Length: 4001
 Score: 2385.00 Matches: 521
 Percent Similarity: 59.98% Conservative: 134
 Best Local Similarity: 47.71% Mismatches: 319
 Query Match: 42.74% Indels: 118
 DB: 9 Gaps: 17

US-10-764-390-3 (1-1072) x AF275679 (1-4001)

Qy	4	ProThrGlyValLeuSerSerLeuLeuLeu-----LeuValThrIleAlaGlyCys-Al	21
Db	297	CTGTACCTGTTTATCTACTTTTGTCTTTCAGCGTCTGTGTGTCAACAGATGCCAG	356
Qy	21	aArgLysGlnCysSerGluGlyArgThrTySerAsnAlaValIleSerProAsnLeuG1	41
Db	357	TCAGAGCAGGTGCCAGCAGGGGAACACAAATTGGAGTTGGCCTGAGATCTGGGGGAGA	416
Qy	41	uThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAlaAlaCy	61
Db	417	A---AATCACTCTGGCTTCTTGAAGGAACCCCTCTCTCCAGTCATGTTGGGCTGCCTG	473
Qy	61	sCysAspLeuSerSerCysAspLeuAlaTrpPheGluGlyArgCysTyLeuValSe	81
Db	474	CTGCCAGGACTGCTGCTGCCATGCTTTTGGTGGCTAGAGGATGTGCATTCCAGGCAGA	533
Qy	81	rCysProHisLysGluAsnCysGluProLysLysMetGlyProIleArgSerTyLeuTh	101
Db	534	CTGCAGCAGGCCCGCAGAGCTGCGGGCTTTTGGAGCACACTCTCCCAATTCCATGCTGT	593
Qy	101	rPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyArgLysMetMetLe	121
Db	594	GTTT---TTAAAAAATTCCAACTGCAGATGTTG-----	627
Qy	121	uAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAspIleArgLysAspLe	141
Db	628	-----GGCTTCTACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG	659
Qy	141	u***PheLeuGlyLysAspTrpGlyLeuGluMetSerGluTyTy***AspAspTyAr	161
Db	660	TCTGGGGCTAGGTTGGAACTGGGCA-----	684
Qy	161	gGluLeuGluLysAspLeuLeuProSerGlyLysGlnGluProArgGlySerAlaG1	181
Db	685	-----TCTTGGAGGAGAGGCCCAACCCAGAGCTGCACCT	716
Qy	181	uTyThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGlyAs	201
Db	717	C-----	717
Qy	201	pSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyLeuAsnGluSe	221
Db	718	-AGACCTGTGTATCTTCCAGTGACCAGCAG-----	747
Qy	221	rAlaSerThrProAlaProLysLeuProGluArg-----SerValLeuLeuPr	237
Db	748	-----AGCTTAATCAGGAAGCTTCAGAGAGAGAGGTAGTCCAGTGCAGTAGTTACACC	800
Qy	237	oLeuThrThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuG1	257
Db	801	TATAGTGACACAGTATCT-----AAAGTGAATGACTCCACGAATTAGGTGGTCTGAC	854
Qy	257	nGluGlnSerSerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProPr	277
Db	855	TACCAGTGGCTCTGCAGAGGTCCACAAGGCGATTACAATTTCAGTCCCTTAACACACAGA	914

Qy	277	oAlaSerLeuLeuLeuSer-----SerValThrValGluLysSerPr	291
Db	915	CTTGACTGCAGAGCTGTGTGGGGCCAAAGATGTATCAGTGCACCTGAAATATCAGA	974
Qy	291	oValLeuThrValThrProGlySer-----ThrGluHisSe	303
Db	975	GGGTCTTGTACTACGCCCACTCAACAGTAAAGATTCTGAGAAAACCCAGATTGC	1034
Qy	303	rIleProThrPro-----ProThrSerAlaAlaProSe	314
Db	1035	TGTCCCCCAGCAGTGGCTCCTCTCAGATTATGCTACCCCTACCCCCAGGCTCTTT	1094
Qy	314	rGluSerThrProSerGluLeuProIleSerProThrThrAlaProArgThrValLysG1	334
Db	1095	CCAGAGCACCTCAGACCATACCCAGTT-----ATAAGGA	1130
Qy	334	uLeuThrValSerAlaGlyAspAsnLeuIleIleThrLeuProAspAsnGluValGluLe	354
Db	1131	ACTGGTGATATCTGTGGAGAGAGTGTCCAGATAACCTGCTTAAGAATGAAGTTCAATT	1190
Qy	354	uLysAlaPheValAlaProAlaProValGluThrThrTyThrTyAsnTyArgLysTrpAsnLe	374
Db	1191	AAATGCATATCTTCCAAAGAACCCCTAAAGAGGAGAAACCTACACCTACGACTGCGAGCT	1250
Qy	374	uIleSerHisProThrAspTyArgLysGlnGlyLysGlnGlyHisLysGlnThrLeuAs	394
Db	1251	GATTACTCATCTAGAGACTACAGTGGAGAAATGGAAGGGAACATATCCAGATCTCTCAA	1310
Qy	394	nLeuSerGlnLeuSerValGlyLeuTyValPheLysValThrValSerSerGluAsnAl	414
Db	1311	ACTATCGAAGCTCACTCCAGGCTGTATGAATTCAAAGTGATTTGAGAGGGTCAAAATGC	1370
Qy	414	aPheGlyGluGlyPheValAsnValThrValLysProAlaArgValAsnLeuProPr	434
Db	1371	CCATGGGGAAGGCTATGTGAACGTGACAGTCAAGCCAGAGCCCCGTAAAGATCGGCCCCC	1430
Qy	434	oValAlaValValSerProGlnLeuGlnLeuThrLeuProLeuThrSerAlaLeuI1	454
Db	1431	CATTGCTATTGTGTCACCTCAGTCCAGGAGATCTCTTTGCCAACCACTTCTACAGTCAT	1490
Qy	454	aAspGlySerGlnSerThrAspAspThrGluIleValSerTyHisTrpGluGluLeAs	474
Db	1491	TGATGGCAGTCAAAAGCACTGATGATGAATAAATCGTTTCAGTACCATCTGGGAGAACTAA	1550
Qy	474	nGlyProPheIleGluGlyLysThrSerValAspSerProValLeuArgLeuSerAsnLe	494
Db	1551	GGGGCTCTTAAGAGAAGAGAGATTCTGAAGATACAGCCATATTAAACCTAAGTAAACT	1610
Qy	494	uAspProGlyAsnTySerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSe	514
Db	1611	CGTCCCTGGGAACCTACCTTTTCAGCTTGTAGTACTGTAGTACTGTAGTACTGTAGTACT	1670
Qy	514	rThrThrAlaAlaLeuIleValAsnAsnAlaValAspTyProProValAlaAsnAlaG1	534
Db	1671	TACTACTGCAACCTCAGACGTGAACAAAGCTGTGGATTACCCCTCTGTGGCCCAAGCAGG	1730
Qy	534	yProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSe	554
Db	1731	CCCCAACCAAGTGAATCACCTCGCCCAAACTCCACCCCTCTTTGGGAACCAAGACAC	1790
Qy	554	rAspAspHisGlnIleValLeuTyTrpSerLeuGlyProGlySerGluGlyLysHis	574
Db	1791	TGATGATCATGGCATCACCAGCTATGAGTGGTCACTCAGCCCAACAGCAGCAAGGAAAGT	1850
Qy	574	sValValMetGlnGlyValGlnThrProTyLeuHisLeuSerAlaMetGlnGluGlyAs	594
Db	1851	CGTGAGATCGAGGGTGTAGAACCAACCACTTACAGCTCTCTGCGATGCAAGAAAGAGA	1910
Qy	594	pTyThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla***ValTh	614
Db	1911	CTACACTTACCAGCTCACAGTGAATGACACATAGACAGCAGGCCACTGCTCAAGTGAC	1970

```
Qy 614 rValIleValGlnProGluAenAsnArgProProValAlaValAlaGlyProAspLysG1 634
Db 1971 CGTTATTGTGCAACTGAAACATTAAGACCTCTCCAGGCAGATCGAGCCCGAGATAAGA 2030
Qy 634 uLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerAspAspHisG1 654
Db 2031 GCTGACCTTCTCTGGATAGACACACCTGGATGGCAGCAAGAGCTCAGATGATCAGAA 2090
Qy 654 yIleValPheThrHisTrpGlu-HisValArgGlyProSerAlaValGluMetGluAsnI 674
Db 2091 AATTATCTCATATCTCTGGGAAAAAACAACAGGACCTGATGGGTGCGAGTCGAGAATG 2150
Qy 674 leAspLysAlaIleAlaThrThrGlyLeuGlnValGlyThrTyHisPheArgLeuT 694
Db 2151 CTAAACAGAGTGTGCTACTGTGCTGCTGCAAGTGGGACCTATGTGTTCACCTTGA 2210
Qy 694 hrValLysAspGlnGlnGlyLeuSerSerThrThrLeuThrValAlaValLysLysG 714
Db 2211 CTGTCAAGATGAGAGGAACTGCAAGCCAGAGCTCTGTGAATGTCATTGTCAAAGAAG 2270
Qy 714 luAenAsnSerProProArgAlaAlaGlyArgHisValLeuValLeuProAsnA 734
Db 2271 AAATAAACAACACCTATAGCAAGATACTGGGAATGTGTGATTACCTACCCACGA 2330
Qy 734 snSerIleThrLeuAspGlySerArgSerThrAspGlnArgIleValSerTyrlLeuT 754
Db 2331 GCACAGCAGACTGATGGCTCTAAGTCTCTCAGATGACAGGAATAGTCAGCTACCTCT 2390
Qy 754 rpIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerV 774
Db 2391 GGACTCGAGATGAGGGAGCCAGCAGCAGGGAGGTGTAAATCACTCTGACCATCACC 2450
Qy 774 alAlaLeuGlnLeuThrAsnLeuValGluGlyValTyThrPheHisLeuArgValThrA 794
Db 2451 CTATCTCTTTTCTTCAAACTGTGTGAGGAACCTACACTTTTCACTGAAAGTGCAG 2510
Qy 794 spSerGlnGlyValSerAspThrAspThrAlaThrValGluValGlnProAspProArgL 814
Db 2511 ATGCAAGGTTGAGAGTGACACAGCCGACCATCTGTGAGGTGAACCTGATCCACGA 2570
Qy 814 ysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgL 834
Db 2571 AAACAACCTGTGTGAGATCATCTTGGATATCAAGTCAGTCAGTCACTGAGAGGTGA 2630
Qy 834 ysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysV 854
Db 2631 AGGGGATTTTCATCGCCAGATTTGGGTCTCTCTGGGGTCTGCTGATTCGCACATCATG 2690
Qy 854 alGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTyrrValGlnSerA 874
Db 2691 TGCAAAAGATTTCAGCCGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2750
Qy 874 rgProProPheLysValLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLeus 894
Db 2751 AGCTTCCCAACAGATCTTCAAGCCATGAGGTGGCAGCGATGCTCAAGAGTGAAGTGC 2810
Qy 894 erLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysL 914
Db 2811 GGAAGCAAAAGCAGACATTTTGTATATTCAGAGCCTTGGAGTCAACACTGTCACATGTC 2870
Qy 914 euLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCysIleCysSerH 934
Db 2871 AGCTGAACCTGTCGACCATGCTGACTCGTTCACCAAAACGCTGTATCTGTGACC 2930
Qy 934 isLeuTrpMetGluAsnLeuIleGlnArgTyrlleTrpAspGlyGluSerAsnCysGluT 954
Db 2931 CTTTTTGTAGAGATTTTCATCAAGGTGCGAGTGGAGGATGGAGACAGCAACTGTGAGT 2990
Qy 954 rpSerIlePheTyrrValThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheT 974
Db 2991 GGAGCGTGTATATGTTATCATCTGCTACCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 3050
Qy 974 hrTrpLeuCysIleCysCysLysArgGlnLysArgThrLysIleArgLysLysThrL 994
```

```
Db 3051 CTGAGCTGTGATCTGTGTGTAAGAGGCAAAA---GGAAAAACCCAAAGGAAAGCA 3107
Qy 994 yStyTrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro-----L 1012
Db 3108 AGTACAAAGATCTGTGATGCCAGGAT--CAGGAAAGCCTGGAGCTGAAGCAACCTCCC 3164
Qy 1012 yStyTrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGluSerG 1032
Db 3165 GAGCAGCATCAAAACAGAAAGCCCTTTTGTGTAAGTAGCAGCCTGATGCATCCGAGTCAG 3224
Qy 1032 luPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnProL 1052
Db 3225 AGCTGGACAGGAC--GATGCCATCTTTACATGGCCAGACCCGAGAGAGGCGCAACTCC 3281
Qy 1052 ysValSerMetAsnGlySerIleArgAenGly 1062
Db 3282 TGCATGTGCAGATGGCTCTGTACCAACGGG 3313

RESULT 9
CQ727300 2970 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION Sequence 13234 from Patent WO02068579.
ACCESSION CQ727300
VERSION CQ727300.1 GI:42292890
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
AUTHORS Kits, such as nucleic acid arrays, comprising a majority of
TITLE humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 13234 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
location/Qualifiers
1..2970
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 2.88e-113 Length: 2970
Pred. No.: 2269.00 Matches: 473
Score: 65.77% Conservative: 115
Percent Similarity: 52.91% Mismatches: 244
Best Local Similarity: 40.66% Indels: 62
Query Match: 6
DB: 12
US-10-764-390-3 (1-1072) x CQ727300 (1-2970)
Qy 218 LeuAsnGluSerAlaSerThrProAlaProLysLeuProGluArg-----Ser 233
Db 283 GTCAACAGTGCACAGCAGAGCTTAATCAGGAAGCTTCAGAAAGAGAGGTAGTCCAGTGAC 342
Qy 234 ValLeuLeuProLeuProThrThrProSerSerGlyGluValLeuGluLysGluLysAla 253
Db 343 GTAGTTACACTATAGTGACACACACTTCT-----AAAGTGAATGACTCAACGAATTA 396
Qy 254 SerGlnLeuGlnGlnSerSerAsnSerSerGlyLysGluValLeuMetProSerHis 273
Db 397 GGTGGTCTGACTACAGTGGCTCTGCAGAGTCCACAGGCGATTACAATTTCCAGTCCC 456
Qy 274 SerLeuProProAlaSerLeuGluLeuSer-----SerValThrVal 287
Db 457 CTAACCCAGACCTGACTGCAGAGCTGTCTGGTGGGCCAAAGATGTATCATGTCACACT 516
Qy 288 GlnLysSerProValLeuThrValThrProGlySer----- 299
Db 517 GAAATATCAGAGGTCTTGTACTACGCCCGAGCTCAACAAAGTAAAGTTCTGAGAAA 576
```



```

Db      2710 AGGAAAGCAAGTACAGATCTCTGGATGCCACGGAT---CAGAAAGCCTGGAGCTGAAG 2766
Qy      1011 Pro-----LysTyrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMetVal 1028
Db      2767 CCAACCTCCGAGCAGGCATCAACAGAAAGGCTTTTGTCTAGTAGCAGCCTGATGCAC 2826
Qy      1029 SerGluSerGluPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArg 1048
Db      2827 TCCGAGTCAGAGCTGGACAGCAT---GATGCCATCTTTTATGTCGCCAGCCGAGAGAAG 2883
Qy      1049 GlyAsnProLysValSerMetAsnGlySerIleArgAsnGly 1062
Db      2884 GGCAAACTCTGATGTCAGATGGCTCTGTATCCCAACGGG 2925

RESULT 10
LOCUS   BD160555          2796 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
ACCESSION
BD160555
VERSION
BD160555.1 GI:27866313
KEYWORDS
JP 2002191363-A/15398.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 2796)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saiko,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE
Primer for synthesizing full-length cDNA and use thereof
JOURNAL
Patent: JP 2002191363-A 15398 09-JUL-2002;
HELIIX RESEARCH INSTITUTE
OS      Homo sapiens (human)
PN      JP 2002191363-A/15398
PD      09-JUL-2002
PF      28-JUL-2000 JP 2000280990
PI      TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
        SAITO,
        JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
        PI      KEIICHI NAGAI, TETSUJI OTSUKI
        PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC      C12P21/02, C12Q1/68/C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT      CDS
source  1. .2796
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      9.74e-109      Length:      2796
Score:          2184.50      Matches:     427
Percent Similarity: 73.14%      Conservative: 93
Best Local Similarity: 60.06%      Mismatches: 186
Query Match:      39.15%      Indels:      5
DB:              6              Gaps:        4

US-10-764-390-3 (1-1072) x BD160555 (1-2796)

Qy      354 LeuLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTyrGluThrAsn 373
Db      3 TTAATGCATATGTTCTCCAGAACCCACCTAAAGGAGAAACCTACACCTACGACTGGCAG 62
Qy      374 LeuLysSerHisProThrAspTyrGlnGlyGluLysGlnGlyHisLysGlnThrLeu 393
Db      63 CTGATTACTATCTCTAGAGCTACAGTGGAGAAATGGAGGAAACATTCACGATCTC 122
Qy      394 AsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAsn 413

```

```

Db      123 AAACATATCGAAGCTCACTCCAGGCTGTATGAATTTCAAAGTGATTGTAGAGGGTCAAAAT 182
Qy      414 AlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgArgValAsnLeuPro 433
Db      183 GCCATGGGGAGGCTATGTGAACGTGACAGTCAAGCCAGAGCCCGCTAAGAATCGGCC 242
Qy      434 ProValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeu 453
Db      243 CCATTGCTATTGTGTACCTCAGTTTCCAGGAGATCTCTTTGGCAACACCTTCTACAGTC 302
Qy      454 IleAspGlySerGlnSerThrAspAspThrGluLeuValSerTyrHisTyrGluGluLeu 473
Db      303 ATTATGGCAGTCAAGCACTGATGATGAATAATCTTCAGTACCAATTGGGAAGAATCTT 362
Qy      474 AsnGlyProPheIleGluGluLysThrSerValAspSerProValLeuLeuLeuSerAsn 493
Db      363 AAGGGGCTCTAAGAGAGAGAGATTTCTGAAGATACAGCCATATTAAACCTAAGTAAA 422
Qy      494 LeuAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsn 513
Db      423 CTCGTCCTCGGAACTACACTTTTCAGCTTGACTGTAGTACTCTGATGGAGCTACCAAC 482
Qy      514 SerThrThrAlaAlaLeuIleValAsnAsnAlaValAspTyrProValAlaAsnAla 533
Db      483 TCTACTACTGCAACCTGACAGTGAACAAGCTGTGATTTACCCCTCTGTGGCCACGCA 542
Qy      534 GlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlnGlnSer 553
Db      543 GGCCCAACCAAGTGTATCACCCTGCCCAAACTCCATCACCCTCTTTGGGAACACAGAGC 602
Qy      554 SerAspAspHisGlnIleValLeuTyrGluTyrSerLeuGlyProGlySerGluGlyLys 573
Db      603 ACTGATGATCATGGCATCACCAGCTATGAGTGGTCACTCAGCCCAAGCAGCAAGGGGAAA 662
Qy      574 HisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGly 593
Db      663 GTGTGGAGATGACAGGCTGTAGAACCAACCTTACAGCTCTCTGCGATGCAAGAGA 722
Qy      594 AspTyrThrPheGlnLeuLysValThrAspSerArgGlnGlnSerThrAla***Val 613
Db      723 GACTACACTTACCAGCTCAGTGTGACATAGACACATAGGACAGCCAGCTGCTCAAGTG 782
Qy      614 ThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaGlyProAspLys 633
Db      783 ACTGTTATTGTCAACCTGAAACCAATAAGCTCTCTCAGGCAGATGCGAGGCCAGATAA 842
Qy      634 GluLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSerSerAspHis 653
Db      843 GAGCTGACCTTCTCTGGATAGCACCAACCTGGATGGCAGCAAGAGCTCAGATGATCAG 902
Qy      654 GlyIleValPheTyrHisTyrGluHisValArgGlyProSerAlaValGluMetGluAsn 673
Db      903 AAATATATCATATCTCTGGGAAAAAACACAGAGGACCTGTATGGGGTGCAGCTCGAGAT 962
Qy      674 IleAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeu 693
Db      963 GCTAACACAGCAGTGTGTCTACTGTGCTGGGCTGCAAGTGGGACCTATGTGTTCACCTTG 1022
Qy      694 ThrValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAlaValLysLys 713
Db      1023 ACTGTCAAAGATGAGAGAACCTCTGCAAGGCCAGAGCTCTGTGAATGTCTTGTCAAAGAA 1082
Qy      714 GluAsnAsnSerProProArgAlaArgAlaGlyArgHisValLeuValLeuProAsn 733
Db      1083 GAATAAACAAACACCTATAGCCAGATACTGGGAATGTGGTATGTACCTACCCACG 1142
Qy      734 AsnSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrLeu 753
Db      1143 AGCACAGCAGAGCTGGATGGCTCTTAAGTCTCAGATGACAGAGGAATAGTACGCTACCTC 1202
Qy      754 TrpIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSer 773

```

Db 1203 TGGACTCGAGATGAGGGAGCCAGCAGCGGGAGGTGTTAAATCACTCTGACCATCAC 1262
Qy 774 ValAlaLeuLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThr 793
Db 1263 CCTATCCTTTCTTCAAACTGGTGGAGAACCTACATTTTTCACCTGAAGTGACC 1322
Qy 794 AspSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArg 813
Db 1323 GATCAAAAGGGTGAGAGTGACACAGCCGACCATGTGGAGGTGAACCTGATCCAGG 1382
Qy 814 LysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlnLeuThrGluGlnArg 833
Db 1383 AAAACAACCTGGTGAGATCATCTTGGATTAACAACCTGAGTCACTTAACGAGGGCTG 1442
Qy 834 LysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLys 853
Db 1443 AHGGGATGTTTCATCCGAGATGGGGTCTCTCTGGGGTGCTGGATTCGACATCATTT 1502
Qy 854 ValGlnLysIleArgAlaHisSerLeuSerThrValIleValPheTyrValGlnSer 873
Db 1503 GTGCAAAAGATTACGCGGTACAGCGAGCAGACCAAAATGGTATTTTGTTCAAAAC 1562
Qy 874 ArgProProPheLysValLeuLysAlaGluValAlaArgAsnLeuHisMetArgLeu 893
Db 1563 GAGCCTCCCCACAGATCTTCAAGGCCATGAGGTGGCAGCGATGCTCAAGAGTGAGCTG 1622
Qy 894 SerLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCys 913
Db 1623 CGAAGCAAAAGGAGACATTTTGATATTTCAGAGCTTGGAAAGTCAACACTGTCACTGT 1682
Qy 914 LeuLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCysIleCysSer 933
Db 1683 CAGCTGAACATGTTCCGACCATCGCCACTGTGACTGTTCCACCAACGCTGATCTGTGAC 1742
Qy 934 HisLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyGluSerAsnCysGlu 953
Db 1743 CCTTTTGGATGGAGAAATTTCAATCAAGGTGCAGCTGAGGGATGGAGACAGCAACTGTGAG 1802
Qy 954 TrpSerIlePheTyrValThrValLeuAlaPheThrLeuIleValLeuThrGlyPhe 973
Db 1803 TGGAGCGTGTATATGTTATCATCTGCTTGTGCTTGTGCTTGTGCTTGGGAATCTCTG 1862
Qy 974 ThrTrpLeuCysIleCysCysLysArgGlnLysArgThrLysIleArgLysLysThr 993
Db 1863 TCTTGGACGTGATCTGTTGTTGAAGAGCAAAA---GGAAACCAAGAGGAAAGC 1919
Qy 994 LysTyrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro----- 1011
Db 1920 AAGTACAGATCCTGGATGCCAGGAT---CAGGAAAGCCTGGAGCTGAAGCCAACTCC 1976
Qy 1012 LysTyrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGluSer 1031
Db 1977 CGAGCAGGATCAACAGAAAGCCCTTTTGTAAAGTAGCGCTGATGCACTCCGAGTCA 2036
Qy 1032 GluPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnPro 1051
Db 2037 GAGTGGACAGCGAT---GATGCCATCTTTTACATGCCACAGCCAGAGAGAGGCAAACTC 2093
Qy 1052 LysValSerMetAsnGlySerIleArgAsnGly 1062
Db 2094 CTGCATGGTGCAGAAATGGCTCTGTACCAACGGG 2126

RESULT 11

AX883831
LOCUS AX883831
DEFINITION Sequence 18736 from Patent EP1074617.
ACCESSION AX883831
VERSION AX883831.1 GI:40038732
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Iehii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 18736 07-FEB-2001;
RESEARCH Association for Biotechnology (JP)
Location/Qualifiers
FEATURES
source
1. 2796
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
96..2171
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE92130.1"
/db_xref="GI:40038733"
/translation="MEGKHSQILKLKLTPLGLYEFKVIVEGQNAHGEYVNVTVKPEP
RKNPPIAVSPQFEISLPTTSVIDGSQDDDKIVQYHWEELKGERBEKISED
AILKSLKLVGNVTPSLTVSDGATNSTANLTNKAVDYPPVANAGPNQVITLUPQ
SITLFGNSDHDGITSYMSLSPSSKGVEMQGVRTPTLQLSAMQSDYTYLTVT
DTGQATAOVTIVQPNKPPQADAGPKDLTPVDSTTLDGSKSSDDQKISLYL
EKTOGPQVLENAANSVATVGTGVGTYYVFTLVKDERNLQSSSVNVIKKEEINKP
PIAKTGNVITLPTSTAEIDGSKSSDDKGLVSVLWTRDEGSPAAGEVLNHSDDHPIL
FLSNLVEGTYFFHLKVDSDIIIVQIQIPTYEQSTKVVFFVQNEPPOHFKGHEVAAMLKSE
LRKQADLPFLRALEVNATVQLNCSHDGCHDSFTKRCICDPFWMENFIKVQLRDGDS
NCESVLVIATFVIIVVALLGILSWTIVICCKROKPKPKSKVKILDATDOESLELK
PTSPAGIKQKGLLLSSLSMSESELSDDAIFTWPDREKGLKHQNGSVPNGQTPLK
ARSPREIL"

ORIGIN

Alignment Scores:
Pred. No.: 9.74e-109 Length: 2796
Score: 2184.50 Matches: 427
Percent Similarity: 73.44% Conservative: 93
Best Local Similarity: 60.06% Mismatches: 186
Query Match: 39.15% Indels: 5
DB: 6 Gaps: 4
US-10-764-390-3 (1-1072) x AX883831 (1-2796)

Qy 354 LeuLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTyrGluTrpAsn 373
Db 3 TTAATGTCATATGTTCTCAAGAACACCTTAAAGGAGAAACCTACACCTACGACGGCAG 62
Qy 374 LeuIleSerHisProThrAspTyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeu 393
Db 63 CTGATTACTATCTTAGAGACTACAGTGGAGAAATGGAGGAAACATTTCCCGAGATCTCTC 122
Qy 394 AsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAsn 413
Db 123 AAACATATCGAAGCTCACTCCAGGCTGTATGAATTCAAAGTGATTGTAGAGGTCATAAT 182
Qy 414 AlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgArgValAsnLeuPro 433
Db 183 GCCCATGGGAGGCTATGTGAACGTGACAGTCAAGCCAGAGCCGCCGTAAGAAATCGGCC 242
Qy 434 ProValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeu 453
Db 243 CCCATTGTATTGTTCACCTCAGTTCAGGAGATCTCTTTGCCAACCACTTCTTACAGTC 302
Qy 454 IleAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHisTrpGluGluIle 473
Db 303 ATTGATGTCAGTCAAAAGCACTGATGATATAAATCGTTTCAGTACCATTTGGGAAGACTT 362
Qy 474 AsnGlyProPheIleGluGlyLysThrSerValAspSerProValLeuArgLeuSerAsn 493
Db 363 AAGGGCCCTTAAGAGAGAGAGAAATTTCTGAAGATACAGCCATATTTAAACCTAGATAA 422
Qy 494 LeuAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyValThrAsn 513
Db 423 CTCGTCCCTGGGAACCTACACTTTGAGTGTAGTACTGTAGTACTGTAGTGGAGCTCAAC 482

Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigetani,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Sato,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamaashita,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL
PUBMED
REFERENCE
AUTHORS

Isogai,T., Ota,T., Hayaashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2796)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES

source

1. .2796
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3004051"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP3"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
96. .2171
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14874.1"
/db_xref="GI:10436632"
/translation="MEGRHSQILKSLKTPGLYEPKVIIEGQNAHGEVYVNVTKPEP
RKNRPPIAIVSQPQESILPTSTTSDIGSQDDDKIVQYHWEELKGLPREKISEDT
AILKSLKVPNGNTSLTVDSGATNSTANLVKNAVDYPPVANAGNPVITLPQN
SITLFGNQSTDHGTISYEWSLSPSSKGVBMQGVPTLQSLAMQSDQKIIISLYM
DTIGQQAQTVTVIQQENKPPQADAGPKELTLPVSTTLDGSKSDQKIIISLYM
EXTQGGQVQLENANSSVAITVGLQGVITVTLTVKDERNLQSQSVNVIKREINKP
PIAKTGNVITLPTSTAEIDGSKSDGDDGVIYSVLWTRDEGSPAAGEVLNHSDDHPIL
FLSNVVEGTVFHLKVTDAKESDTRITVEYKPPDKNNLVEIILDINVSQTLERLK
GMFRIQIGLVGLSDIIVQKIQPYQETKMFVQNEPPHQIFKGEHVAAMLKSE
LRKQKADFLIIFRALENTVTTCOLNCSHGCHDSFTKRCICDPFMENFIKVLORDGDS
NCWMSVLYVLIATFVIVAGLISWTVICCCCKQKQKPKSKYKILIDATDESLELK
PISRAGIKQKGLLSLSLHSESELDSDDAITFPDREKGLHQQNGSVPMNGQPLK
ARSPREEL"

CDS

ORIGIN

Alignment Scores:

Pred. No.: 9,74e-109 Length: 2796
Score: 2184.50 Matches: 427
Percent Similarity: 73.14% Conservative: 93
Best Local Similarity: 60.06% Mismatches: 186
Query Match: 39.15% Indels: 5
DB: Gaps: 4

US-10-764-390-3 (1-1072) x AK024287 (1-2796)
QY 354 LeuylsAlaPheValAlaProAlaProValGluThrThrTyrAsnTyrGluTrpAsn 373
Db 3 TTTAAATGCATTTGTTCCTCAAGAAACCTTAAAGAGAGAACCTACACTACGACTGGCAG 62
QY 374 LeuileSerHisProThrAspTyrGlnGlyGluileLysGlnGlyHisLysGlnThrLeu 393
Db 63 CTGATTACTCATCTAGAGACTACAGTGGAGAAATGGAGGGAACATTCCTCCAGATCTC 122
QY 394 AsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAsn 413
Db 123 AAATCTCGAAGCTCACTCCAGGCTGTATGAATTCAAAGTGATTGTAGAGGGTCAAAAT 182
QY 414 AlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgArgValAsnLeuPro 433
Db 183 GCCCATGGGGAAGGCTATGTGAAGTGACAGTCAAGCCAGAGCCGCCGTAAGAATCGGCC 242
QY 434 ProValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeu 453
Db 243 CCCATTGCTATTGTGTCACTCAGTTCAGAGAGATCTCTTGGCCACACCTTCTACAGTC 302
QY 454 IleAspGlySerGlnSerThrAspThrGluileValSerTyrHisTyrGluGluLurle 473
Db 303 ATTGATGCGCAGTCAAGGACCTGATGATAAAATCGTTCACTACCATTTGGGAAGACTT 362
QY 474 AsnGlyProPheIleGluLysThrSerValAspSerProValLeuArgLeuSerAsn 493
Db 363 AAGGGGCTCTTAAGAGAAGAGAGATTTCTGAAAGATACAGCCATATTAATACTAAGTAA 422
QY 494 LeuAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyValAlaThrAsn 513
Db 423 CTGTCCTCGGAACCTACACTTTCACTTGTAGTGTAGTACTCTGATGGAGTACCAAC 482
QY 514 SerThrThrAlaAlaLeuileValAsnAsnAlaValAspTyrProProValAlaAlaAla 533
Db 483 TCTACTACTGCAACCTGACAGTGAACAAAGCTGTGGATTACCCCTCTGTGGCAACGCA 542
QY 534 GlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSer 553
Db 543 GGGCCCAACCAAGTGATCACCCTGCCCCAAACCTCCATCACCTCTTTGGGAACACAGAGC 602
QY 554 SerAspAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLys 573
Db 603 ACTGATGATCATGGCATCACCAGCTATGAGTGGTCTCTCAGCCCAAGCAGCAAGGGA 662
QY 574 HisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGly 593
Db 663 GTGGTGGAGATGCAAGGTGTTAGAACCAACCTTTACAGTCTCTCTGCGATGCAAGAAGA 722
QY 594 AspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla**Val 613
Db 723 GACTACACTTACCAGCTCACAGTCACTGACACAAATAGGACAGAGCCACTCTCTCAAGTG 782
QY 614 ThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaGlyProAspLys 633
Db 783 ACTGTTATGTGCAACCTGAAACAAATAGCTCTCTCAGCCAGATGCGAGGCCCAAGATAA 842
QY 634 GluLeuIlePheProValGlnSerAlaThrLeuAspGlySerSerSerSerAspAspHis 653
Db 843 GAGCTGACCTTCTCTGTGGTAGGACCAACCTTGGATGGAGGACCTGATGATGATCAG 902
QY 654 GlyIleValPheTyrHisTyrGluHisValArgGlyProSerAlaValGluMetGluAsn 673
Db 903 AAAATTATCTCATATCTCTGGGAAAAACACAGGGACCTGTATGGGGTGCAGATCGAGAAT 962
QY 674 IleAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeu 693
Db 963 GCTAAACAGCAGTGTGTACTGTGTGCTGGCTGCAAGTGGGAGCCTATGTGTTCACCTTG 1022
QY 694 ThrValLysAspGlnGlnGlnGlyLeuSerSerThrThrLeuThrValAlaValLysLys 713

/note="synonyms: KIAA1837, PP791"
/db_xref="LocusID:79932"

CDS

<1_2102
/gene="PKD1-like"

/codon_start=3

/product="PKD1-like protein"

/protein_id="AAH31672.2"

/db_xref="GI:48257271"

/db_xref="LocusID:79932"

/translation="HPRDYSMEGKHSQILKSLTGLYEFKVIYVEQNAHGEVY
NTVPEPRKNPPIATVSPQFIEISLFTSTVSDGSDDDKILVTHWELKGPLR
BEKSDTAIILKSLKPGNFTSVVSDGATNSTANITVANKADVPPVANAGEN
QVITLPQNSITLFGNSQDDHGTITSVSPSSKGVEMQGVRTPTLQSAHQEGD
YTYQLTVDTTGOQATQAVTVIQPENKPPQADAPDKELTLPVDSFTLQSGSSDD
OKIISYLEKTKGPGDVOLENANSSVATVTVGCVTVYFTLVKDERNLQSSVNVV
VKEENKPKPIAKITGNVITPTSTAEIDGSKSSDDKGLVSYLWTRDRSGSPAAGEVLA
HSDHPILFNSLVNVEGTFVHLKVDAGESDTRITVEVPDPKRNKLVEIILQINV
QSTRLKGMFTQGVLLGLSDSIIIVOKIQPYTEQTKMVFVQNSPPHPIFKGHE
VAAMLSLRKQADFLIIFRALENTVTCOLNSDHGCHDSFTKRCICDPFWMENFIK
VOLRDGNSCEWSVLYIIATFVIIVVALLISWTVICCKRQKPKRKSFKXILIDAT
DOESLEKPTSPAGIKOKGLLLSSLMHSELSDDDAIFWPDREKGLLHGQNGSV
FNGQPLKARSPEBIL"

ORIGIN

Alignment Scores:

Pred. No.: 2,12e-105 Length: 2743
Score: 2122.50 Matches: 416
Percent Similarity: 73.55% Conservative: 90
Best Local Similarity: 60.47% Mismatches: 177
Query Match: 38.04% Indels: 5
DB: 9 Gaps: 4

US-10-764-390-3 (1-1072) x BC031672 (1-2743)

QY 377 HisProThrAspTyrGlnGlyGluLeuLysGlnGlyHisLysGlnThrLeuAsnLeuSer 396
Db 3 CATCTAGACGTACAGTGGAGAAATGGGAAGGAAACATTCCTCCAGATCTCTAAATATCG 62
QY 397 GlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAsnAlaPheGly 416
Db 63 AAGCTCACTCCAGGCTGTGATGAATTCAGAGGATTTAGAGGGTCAAAATGCCATGGG 122
QY 417 GluGlyPheValAsnValThrValLysProAlaArgValAsnLeuProProValAla 436
Db 123 GAAGGCTATGTGACGTGACAGTCAAGCCAGAGCCCGTGAAGATCGGCCCCCATGCT 182
QY 437 ValValSerProGlnLeuGlnLeuThrLeuProLeuThrSerAlaLeuLeuAspGly 456
Db 183 ATTGTGTCACTCAGTTCAGGAGATCTCTTTGCCAACCACTTCTACAGTCATTGATGGC 242
QY 457 SerGlnSerThrAspAspThrGluLeuValSerTyrHisTrpGluLeuAsnGlyPro 476
Db 243 AGTCAAGACACGTATGATGATAAATCGTTCAGTACCATTGGAGAGAACTTAAGGGCGCT 302
QY 477 PheIleGluLysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspPro 496
Db 303 CTAAGAGAGAGAGAGATTCTCGAGATACAGCCATATTAACCTAAGTAACTCGTCCCT 362
QY 497 GlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThr 516
Db 363 GGGAACTACATCTTCAGCTTGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 422
QY 517 AlaAlaLeuLeuValAsnAsnAlaValAspTyrProProValAlaAsnAlaGlyProAsn 536
Db 423 GCAAACTCAGAGTGAACAAAGCTGTGGATTAACCCCTGTGGCCCAACGAGGCCCAAC 482
QY 537 HisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlnLysGlnSerSerAspAsp 556
Db 483 CAAGTATCACCCTGCCCAAACTCCATCACCTCTTTGGGAAACCAAGAGACTGATGAT 542
QY 557 HisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHisValVal 576
Db 543 CATGGCATCACCAGCTATGAGTGTGCTCACTCAGCCCCAAGCAGCAAGGAAAGTGTGGAG 602

QY 577 MetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAspTyrThr 596
Db 603 ATGAGGGGTGTAGAACCAACCACTTACAGCTCTCTGGATGCAAGAGAGACTACACT 662
QY 597 PheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla***ValThrValle 616
Db 663 TACCAGCTTCAGTGTGACTGACACAAATAGGACAGCAGCCACTGTCTCAAGTACTGTATT 722
QY 617 ValGlnProGluLeuAsnArgProProValAlaValAlaGlyProAspLysGluLeuLeu 636
Db 723 GTGCAACTGTAACAAATAAGACCTCTCAGCAGATGCGGCCAGATTAAGAGCTGACC 782
QY 637 PheProValGlnSerAlaThrLeuAspGlySerSerSerSerAspAspHisGlyIleVal 656
Db 783 CTTCCTGTGGATAGACACCACTCTGATGGCAGCAAGAGCTCAGATGATCAGAAAAATTATC 842
QY 657 PheTyrHisTrpGluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLys 676
Db 843 TCATATCTCTGGGAAAAACACAGGACCTGATGGGTGCGAGCTCGAGATGCTAACACAGC 902
QY 677 AlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrValLys 696
Db 903 AGTGTGTCTACTGTGACTGGGCTGCAAGTGGGACCTATGTGTCCACTTGACTGTCAAA 962
QY 697 AspGlnGlnGlyLeuSerSerThrThrLeuThrValAlaValLysLysGluAsnAsn 716
Db 963 GATGAGAGAACTTCAAAGCCAGAGCTCTGTGAATGTTCATTGTCAAGAGAGAAATAAC 1022
QY 717 SerProProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAsnSerIle 736
Db 1023 AAACCCATATAGCAAGATAAATCTGGGAATGTGGTGTATACCTACCACGAGCACAGCA 1082
QY 737 ThrLeuAspGlySerAspSerThrAspAspGlnArgIleValSerTyrLeuTrpIleArg 756
Db 1083 GAGCTGGATGCTCTAAGTCTCTCAGTACAGAGGAATAGTACACTCTCTGGACTCGA 1142
QY 757 AspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeu 776
Db 1143 GATGAGGGAGCCCAAGCAGCAGAGGGAGGTGTTAAATCACTCTGACCAATCCCTATCCT 1202
QY 777 GlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThrAspSerGln 796
Db 1203 TTTCTTTCAAACTGGTTGAGGAACTTACACTTTTCACTTCACTGAAAGTACCGATCAAG 1262
QY 797 GlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArgLysSerGly 816
Db 1263 GGTGAGGTGACACAGACCGGACCACTGTGGAGGTGAACCTGTATCCAGGAAACCAAC 1322
QY 817 LeuValGluLeuThrLeuGlnValGlyGlnLeuThrGluGlnArgLysAspThr 836
Db 1323 CTGGTGGAGATCATCTTGGATATCAACCTCAGTCAGCTAATCAGAGGCTGAAGGGATG 1382
QY 837 LeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLys 856
Db 1383 TTCTCCGCCAGATTTGGGGTCTCTCTGGGGGTGCTGGATTCGACATCATTTGTGCAAAAG 1442
QY 857 IleArgAlaHisSerAspLeuSerThrValIleValPheTyrValGlnSerArgProPro 876
Db 1443 ATTACGCGTACACGAGCAGACCAACCAAAATGGTATTTTGTTCAAACGAGGCTCCC 1502
QY 877 PheLysValLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLysGlu 896
Db 1503 CACCAGATCTTCAAGGCCATGAGGTGGCAGCGATGCTCAAGAGTGAGCTGGGGAAGCAA 1562
QY 897 LysAlaAspPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLys 916
Db 1563 AAGCAGACTTTTTCATATTCAGAGCTTGGAGCTTGAAGTCAACACTGTGCATGTGAGTGAAC 1622
QY 917 CysSerGlyHisGlyHisCysAspProLeuThrLysArgCysValleCysSerHisLeuTrp 936
Db 1623 TGTTCCGACCATGGCCACTGTGACTCGTTCCAAACCGCTGTATCTGTGACCCCTTTTGG 1682

Qy 937 MetGluAanLeuLeuGlnArgTyIleTrpAspGlyGluSerAsnCysGluTrpSerIle 956
 Db 1693 ATGGAGAATTCATCAAGTGCAGCTGAGGATGAGACAGACAACTGTGAGTGGAGCGTG 1742
 Qy 957 PheTyValThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTripleu 976
 Db 1743 TTATATGTTATCATTTGCTACCTTTGTCATTGTTGCTTGGGAATCCTGCTTGGACT 1802
 Qy 977 CysIleCysCysCysGlyArgGlnLysArgThrLysIleArgLysLysThrLysTyThr 996
 Db 1803 GTGATCTCTTGTGTAAGAGGCAAAA--GGAAAAACCAAGAGAAAGCAAGTACAAG 1859
 Qy 997 IleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro-----LysTyGly 1014
 Db 1860 ATCTGGATGCCAGGAT--CAGAAAGCTGAGGCTGAAGCCAACTCCCGAGCAGGC 1916
 Qy 1015 IleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGluSerGluPheAsp 1034
 Db 1917 ATCAACAGAAAGGCTTTTCTTAAGTAGCAGCCTGATGCACTCCGAGTCAGAGCTGGAC 1976
 Qy 1035 SerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnProLysValSer 1054
 Db 1977 AGCGAT--GATGCCATCTTTACATGCGCAGACCGAGAGAGGGCAAACTCCTGCATGGT 2033
 Qy 1055 MetAsnGlySerIleArgAsnGly 1062
 Db 2034 CAGAATGGCTCTGTACCCCAACGGG 2057

RESULT 14
 LOCUS AB058740 4423 bp mRNA linear PRI 05-JUL-2002
 DEFINITION Homo sapiens mRNA for KIAA1837 protein, partial cds.
 ACCESSION AB058740
 VERSION AB058740.1 Gi:14017890
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R. and Ohara, O.
 Prediction of the coding sequences of unidentified human genes. XX.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro
 DNA Res. 8 (2), 85-95 (2001)
 2 (bases 1 to 4423)
 Ohara, O., Nagase, T. and Kikuno, R.
 Direct Submission
 Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute,
 Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail: cdna@info.kazusa.or.jp)
 URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,
 Fax: 81-438-52-3914

FEATURES
 Location/Qualifiers
 1..4423
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="fh2204081"
 /tissue_type="brain"
 /clone_lib="pYO326GW"
 /note="vector:pBluescriptII SK plus"
 1..4423
 /gene="KIAA1837"
 <28..1947
 /gene="KIAA1837"
 /note="Start codon is not identified"
 /codon_start=1
 /product="KIAA1837 protein"
 /protein_id="BAB47466.1"
 /db_xref="GI:14017891"

gene
 CDS

/translation="VMECKHSQILKLSKLTPLGLYEFKVIYVEGQNAHGEQYVNTVKPE
 PKRNPPIALISFQFIEISLPTSTVIGDSQSTDDDKIVQTHMEELKGPUREEKISED
 TALKLSKLVPGNYTFLSLVDSGATNSTANLTANKVADYPPVANAGQVITLPO
 NSITLFGNQSTDDHGITSYEWLSPSKSKVEMQGVRTPTLQLSAMQEGDYTYQLTV
 TDITGOQATQAVTVIQPENKPPQADAGPKELTLVDSTTLDGSSKSDDKIISYL
 WEKTQPDGVQLENANSSVATVGLQGVTVFTLVKDERNLOSQSSVNVIVKEEINK
 PPIAKITGNVITLPTSTAEIDGSSDDKGIYSYLWTRDEGSPAAGSEVLNHSDDHPI
 LFLSNLVEGYTHLKVTKDAKGSDDTDRRTVEVKPDKRKNLVEIILDINVSQUTERL
 KGMPIRQIGVLLGLVLDSDIIVQIKIPYTESKMFVQNEPPEHQIIPKGEHVAAMLKS
 ELKQRADFLIFRALEVNVTQLCNSDHGHCDSFTKRCICDPFMENFTKVOQRDGD
 SNEWSVLYVIIATFVITVALGILSWTIVCCCKRQKQKPKRKSRYKILDTQDSLEL
 KPTSRAGRGPQCQSF"

ORIGIN
 Alignment Scores:
 Pred. No.: 8e-99 Length: 4423
 Score: 2005.50 Matches: 388
 Percent Similarity: 74.92% Conservative: 84
 Best Local Similarity: 61.59% Mismatches: 155
 Query Match: 35.94% Indels: 3
 DB: 9 Gaps: 3

US-10-764-390-3 (1-1072) x AB058740 (1-4423)

Qy 383 GlyGluLeuLysGlnGly---HisLysGlnThrLeuAsnLeuSerGlnLeuSerValGly 401
 Db 22 GGATGAGTAATGGAAAGGAAACATTCACAGATCTCAAACCTATCGAAGCTCACCTCCAGGC 81
 Qy 402 LeuTyValPheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsn 421
 Db 82 CTGTATGAATTCAAAGTGATTTAGAGGGTCAAAATGCCCATGGGAAGGTATGTGAAC 141
 Qy 422 ValThrValLysProAlaArgValAsnLeuProProValAlaValValSerProGln 441
 Db 142 GTGACAGTCAAGCCAGAGCCCGTAAGAATCGGCCCCCATTTGTTATGTGTACCTCAG 201
 Qy 442 LeuGlnGluLeuThrLeuProLeuThrSerAlaLeuLeuAspGlySerGlnSerThrAsp 461
 Db 202 TTCCAGAGATCTCTTTGGCCAAACCATCTTACAGCTGATGATGCGAGTCAAGCAGCTGAT 261
 Qy 462 AspThrGluLeuValSerTyHisTrpGluGluLeuAsnGlyProPheIleGluGluLys 481
 Db 262 GATGATAAATCGTTCAGTACCATTTGGGAAGAACTTAAGGGGCTCTTAAGAGAGAGAGAG 321
 Qy 482 ThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTySerPhe 501
 Db 322 ATTTCTGAAGATACAGCCATATTAAACTAAGTAACTCGTCCCTGGGAACACTACACTTC 381
 Qy 502 ArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuLeuVal 521
 Db 382 AGCTTGACTGTAGTAGACTCTGATGGAGCTACCAACTCTACTACTGCAAACTGACAGTG 441
 Qy 522 AsnAsnAlaValAspTyProProValAlaAsnAlaGlyProAsnHisThrIleThrLeu 541
 Db 442 AACAAAGCTGTGGATTACCCCTCTGTGCCAACCGAGGCCCAACCAAGTATCACCCTG 501
 Qy 542 ProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspHisGlnIleValLeu 561
 Db 502 CCCCAAACTCCATCACTCTTTGGGAACACAGAGCACTGATGATGATGATGATGATGATGATG 561
 Qy 562 TyrGluTrpSerLeuGlyProGlySerGluGlyLysHisValValMetGlnGlyValGln 581
 Db 562 TATGAGTGTGCTCACTCAGCCCAAGCAGCAAGGGAAGTGGTGGAGATGTCAGGGGTGTAGA 621
 Qy 582 ThrProTyLeuHisLeuSerAlaMetGlnGluGlyAspTyThrPheGlnLeuLysVal 601
 Db 622 ACCAACTTACAGCTCTCTGCGATGCAAGAGAGAGACTACACTTACCAGCTCACAGTG 681
 Qy 602 ThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnProGluAsn 621
 Db 682 ACTGACACATAGGACAGAGCCGACCTGCTCAAGTGAAGTGTATTATGTGCAACCTGAAAC 741
 Qy 622 AsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProValGluSer 641


```
Qy 576 ValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluCysPyr 595
Db 361 GAGATGCAAGGGTGTGAACCAACCACTTACAGCTCTCTGCGATGCAAGAAGGAGACTAC 420
Qy 596 ThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla**ValThrVal 615
Db 421 ACTTACCAGCTCAGAGTCACTGACACATAAGGACAGCAGGCGACTGTCTCAAGTGACTGTT 480
Qy 616 IleValGlnProGluAsnArgProProValAlaValAlaGlyProAspLysGluLeu 635
Db 481 ATTGTGCAACCTGAAACAATAAGCTCTCTCAGGCAGATGCGAGGCCCAAGATAAGAGCTG 540
Qy 636 IlePheProValGlnSerAlaThrLeuAspGlySerSerSerSerSerSerSerAlaHisGlyIle 655
Db 541 ACCCTTCTCTGGATGAGACACACCTCTGATGCGCAAGAGCTCAGATGATCAGAAAAATT 600
Qy 656 ValPheTyrHisTrpGluHisValArgGlyProSerAlaValGluMetGluAsnIleAsp 675
Db 601 ATCTCATATCTCTGGAAAAACACAGGAGCTGTATGGGTGCAGCTCGAGNATCCTAAC 660
Qy 676 LysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrVal 695
Db 661 AGCAGTGTGTACTGTACTGTGGCTGCAAGTGGGACCTATGTGTTCACCTTGACTGTC 720
Qy 696 LysAspGlnGlnGlyLeuSerSerThrThrThrThrValAlaValLysLysGluAsn 715
Db 721 AAGATGAGAGGAACCTTCAAGCCAGAGCTCTGTGAATGTCTTGTCAAGAGAATA- 779
Qy 716 AsnSerProProArgAlaArgAlaGlyArgHisValLeuValLeuProAsnAsnSer 735
Db 780 AACAAACCACTATAGCCCAAGATAAATCGGAATGTGTGATTACCTTACCACGAGCACA 839
Qy 736 IleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrLeuTrpIle 755
Db 840 GCAGAGCTGGATGGCTCTAAGTCTCTCAGATGACAAAGGGAATAGTCAGCTACCTCTGGACT 899
Qy 756 ArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerValAla 775
Db 900 CGAGATAGGGGGAGCCCGAGCAGCAGGGAGGTGTAAATCACTCTGACCATCACCTATC 959
Qy 776 LeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThrAspSer 795
Db 960 CTTTTTCTTCAAACTCTGGTGGAGAACCTACACTTTTTCACCTGAAAGTGACCGATGCA 1019
Qy 796 GlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArgLysSer 815
Db 1020 AAGGCTGAGAGTGACACAGACCGGACCCTGTGGAGGTGAAACCTGTATCCAGGAAAAAC 1079
Qy 816 GlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAsp 835
Db 1080 AACCTGGTGGAGATCATCTTGGATATCAACGCTCAGTCAACTGAGAGGCTGAAGGGG 1139
Qy 836 ThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGln 855
Db 1140 ATGTTTCATCCGCGAGATTGGGGTCTCTCTGGGGGTGCTGGATTCGACATCATTTGTGCAA 1199
Qy 856 LysIleArgAlaHisSerAspLeuSerThrValIleValPheTyrValGlnSerArgPro 875
Db 1200 AAGATTACGCGGTACACGCGCAGAGCACCACAAATGGTATTTTGTTCACAAACGAGCCT 1259
Qy 876 PropheLysValLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLys 895
Db 1260 CCCCACCAAGATTTCAAAAGGCCATGAGGTGGCAGCGATGCTCAAGAGTGAGCTGCGAAG 1319
Qy 896 GluLysAlaAspPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeu 915
Db 1320 CAAAAGGCGAGCTTTTGTATTTACAGCCCTTGGAGTGAACACTGTACATGTACAGCTG 1379
Qy 916 LysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCysIleCysSerHisLeu 935
Db 1380 AACTGTTCCGACCATGGCTACTGTGACTCGTTCACCAACCGCTGTATCTGTGACCCCTTTT 1439
Qy 936 TrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyGluSerAsnCysGluTrpSer 955
```

```
Db 1440 TGGATGAGGATTTTCATCAGGTCCAGCTGAGGATGGAGACACCAACTGTGATGGAGC 1499
Qy 956 IlePheTyrValThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTrp 975
Db 1500 GTGTTATATGTTATCATTTGCTTACCTTTGTTCATTTGTTGCTTGGGAATCCTGTCTTGG 1559
Qy 976 LeuCysIleCysCysCysLeuArgGlnLysArgThrLysIleArgLysLeuThrLysTyr 995
Db 1560 ACTGTGATCTGTTGTTGTAAGAGCAAAAA---GGAAAAACCCAAAGAGAAAGCAAGTAC 1616
Qy 996 ThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro-----LysTyr 1013
Db 1617 AAGAT-CTGGATGCCAGGAT---CAGGAAAGCTTGGAGCTGAAGCCCAACCTCCCGACA 1672
Qy 1014 GlyIleLysHisArgSerThrGluHisAsnSerSerSerLeuMetValSerGluSerGluPhe 1033
Db 1673 GGCATCAACACAGAAAGGCCCTTTTCTAAGTAGCAGCCTGTATGCATCTCCGAGTCAGAGCTG 1732
Qy 1034 AspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnProLysVal 1053
Db 1733 GACAGCGAT---GATGCCATCTTTTACATGGCCAGACCGAGAGAGGGCAAACTCCTCTGCAT 1789
Qy 1054 SerMetAsnGlySerIleArgAsnGly 1062
Db 1790 GGTCAAGATGGCTCTGTACCAACGGG 1816
```

Search completed: October 12, 2005, 13:41:44
Job time : 10126 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:00:24 ; Search time 125 Seconds
(without alignments)
3316.857 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 5580

Sequence: 1 MAPPTGVLSLLLVTTIAGC.....VSMNGSIRNGASFYSCKSR 1072

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1980s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5576	99.9	1072	7 ADJ69816	Adj69816 Human hea
2	5576	99.9	1072	8 ADR00600	Adr00600 Human 254
3	5574	99.9	1072	8 ADR00601	Adr00601 Human 254
4	5574	99.9	1072	8 ADR00597	Adr00597 Human 254
5	5574	99.9	1072	8 ADR00598	Adr00598 Human 254
6	5574	99.9	1072	8 ADR00848	Adr00848 Human 254
7	5574	99.9	1072	8 ADR00594	Adr00594 Human 254
8	5574	99.9	1072	8 ADR00592	Adr00592 Human 254
9	5485	98.3	1053	8 ADR00862	Adr00862 Human 254
10	5485	98.3	1053	8 ADR00863	Adr00863 Human 254
11	5485	98.3	1063	8 ADR00861	Adr00861 Human 254
12	5485	98.3	1063	8 ADR00596	Adr00596 Human 254
13	5485	98.3	1063	8 ADR00599	Adr00599 Human 254
14	4938	88.5	978	7 ADG14994	Adg14994 Human SEC
15	2426.5	43.5	1048	8 ADM36227	Adm36227 Murine SS
16	2401	43.0	1049	4 AAE05768	Aae05768 Human sec
17	2401	43.0	1049	7 ADI21080	Adi21080 Novel hum
18	2086.5	37.4	691	4 AAB95784	Aab95784 Human pro
19	1431.5	26.0	1069	4 ABB63364	Abb63364 Drosophil
20	1440.5	25.8	685	4 AAB94231	Aab94231 Human pro
21	1375	24.6	711	4 ABG08994	Abg08994 Novel hum
22	1296.5	23.2	640	7 ADP42454	Adp42454 Human pp7
23	823	14.7	272	6 ABU70371	Abu70371 Human adi
24	815.5	14.6	346	4 AAM00902	Aam00902 Human bon
25	487	8.7	162	4 AAO09933	Aao09933 Human pol

ALIGNMENTS

RESULT 1

ADJ69816

ID ADJ69816 standard; protein; 1072 AA.

XX AC ADJ69816;

DT 06-MAY-2004 (first entry)

DE Human heat mitochondrial protein as a therapeutic target SeqID1622.

XX KW Mitochondrial; human; screening assay; diabetes mellitus;

XX KW Hunting's disease; osteoarthritis;

XX KW Leber's hereditary optic neuropathy; LHON;

XX KW Mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

XX KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

XX KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;

XX KW osteopathic; ophthalmological; cytostatic.

OS Homo sapiens.

XX WO2003087768-A2.

PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010870.

XX PR 12-APR-2002; 2002US-0372843P.

PR 17-JUN-2002; 2002US-0389987P.

PR 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.

PA (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Pahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

PI Warnock DE;

XX WPI; 2003-845369/78.

DR Identifying a mitochondrial target for drug screening assays and for

XX treating diseases associated with altered mitochondrial function,

PT comprises detecting a modified polypeptide in a sample and correlating

PT with the disease.

XX Claim 1; SEQ ID NO 1622; 180pp; English.

PS This invention relates to novel mitochondrial targets that can be used

XX for therapeutic intervention in treating a disease associated with

CC altered mitochondrial function. Specifically, it refers to a method for

CC altered mitochondrial function.

CC altered mitochondrial function.

CC altered mitochondrial function.

CC altered mitochondrial function.

CC altered mitochondrial function.

CC altered mitochondrial function.

CC altered mitochondrial function.

CC altered mitochondrial function.

identifying proteins of the human heart mitochondrial proteome that are useful for drug screening assays, as well as therapeutic targets. The present invention describes a method for identifying such proteins that can be used in the treatment of various diseases associated with altered mitochondrial function including diabetes mellitus, Huntington's disease, osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy ragged red fibre syndrome (MERRF) or cancer. Accordingly, these compositions have neuroprotective, neurotropic, antidiabetic, anticonvulsant, antiarthritic, osteopathic, ophthalmological and cytostatic activities. This polypeptide sequence is a human heart mitochondrial protein of the invention.

XX Sequence 1072 AA;

Query Match 99.9%; Score 5576; DB 7; Length 1072;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPPTGVLSLLIIITIAACARKQCEGRYSNAVISPNLETRIMRVSHTFPVVDTAA 60
 DB 1 MAPPTGVLSLLIIITIAACARKQCEGRYSNAVISPNLETRIMRVSHTFPVVDTAA 60
 QY 61 CDDLSLCLAWFEGRCYLVSCPHKENCEPKMGPIRSYLTFLVLPVORPAQLLDYDMM 120
 DB 61 CDDLSLCLAWFEGRCYLVSCPHKENCEPKMGPIRSYLTFLVLPVORPAQLLDYDMM 120
 QY 121 LNRGSPGSGIDSPDIRKDLFLGKDWGLEMSYDDYREKDLQLQSGKQPRGSA 180
 DB 121 LNRGSPGSGIDSPDIRKDLFLGKDWGLEMSYDDYREKDLQLQSGKQPRGSA 180
 QY 181 EYTDWGLLPGEAGANSSVGDSPVPAETQDPPELHYLNESASTPAPKLPERSVLLPPT 240
 DB 181 EYTDWGLLPGEAGANSSVGDSPVPAETQDPPELHYLNESASTPAPKLPERSVLLPPT 240
 QY 241 TFSSEVLEKEKASQLQSGSSNSKEVLMPSHSLPPASLELSSVTVEKSPVLVTTPGST 300
 DB 241 TFSSEVLEKEKASQLQSGSSNSKEVLMPSHSLPPASLELSSVTVEKSPVLVTTPGST 300
 QY 301 EHSIPTPTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
 DB 301 EHSIPTPTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
 QY 361 APPVETTYNENLISHPTDYQGEIKQGHKQTLNLSQLSVGLYVPKVTVSSENARFGEFV 420
 DB 361 APPVETTYNENLISHPTDYQGEIKQGHKQTLNLSQLSVGLYVPKVTVSSENARFGEFV 420
 QY 421 NVTVKPARRVNLPPVAVVSPQLETLPLTSALIDGQSTDDTEIVSVHWEELNPFTEE 480
 DB 421 NVTVKPARRVNLPPVAVVSPQLETLPLTSALIDGQSTDDTEIVSVHWEELNPFTEE 480
 QY 481 KTSVDSPLVRLSNLDPGNSYFRLTVTSDGATNSTTAALIVNNAVDPYPPVANAGPNHIT 540
 DB 481 KTSVDSPLVRLSNLDPGNSYFRLTVTSDGATNSTTAALIVNNAVDPYPPVANAGPNHIT 540
 QY 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGEGKHVVMQGVPPYHLHSAMQEGDYTFOLK 600
 DB 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGEGKHVVMQGVPPYHLHSAMQEGDYTFOLK 600
 QY 601 VTDSRQOSTAXVTIVOPENNRPPVAVAGDPKELIPVESATLDGSSDDHGIVFVHW 660
 DB 601 VTDSRQOSTAXVTIVOPENNRPPVAVAGDPKELIPVESATLDGSSDDHGIVFVHW 660
 QY 661 EHVRFPSAVEMENIDKATATVGLQVGFHFLTVKDOQGLSSTSTLTIVAVKKNNSPPR 720
 DB 661 EHVRFPSAVEMENIDKATATVGLQVGFHFLTVKDOQGLSSTSTLTIVAVKKNNSPPR 720
 QY 721 ARAGGRHVLVLPNNSITLDGSRSTDDQRTVSVLWIRDGQSPAGDVIDGSDHVSVALQLTN 780
 DB 721 ARAGGRHVLVLPNNSITLDGSRSTDDQRTVSVLWIRDGQSPAGDVIDGSDHVSVALQLTN 780
 QY 781 LVEGVYTHLRVTDGASDITATVEQVDPDRKSGLVLTQLQGVGQLTEQRKDTLVRQ 840

Db 781 LVEGVYTHLRVTDGASDITATVEQVDPDRKSGLVLTQLQGVGQLTEQRKDTLVRQ 840
 QY 841 LAVLLNVLDSIDIKVQKIRAHSDLSITVIVFYVQSPFPFKVLAEEVARNLHRLSKEKADF 900
 Db 841 LAVLLNVLDSIDIKVQKIRAHSDLSITVIVFYVQSPFPFKVLAEEVARNLHRLSKEKADF 900
 QY 901 LLFKVLRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYTIWDGESNCEWSIFYVT 960
 Db 901 LLFKVLRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYTIWDGESNCEWSIFYVT 960
 QY 961 VLAFTLIVLTGFTWLTCICCCCKRQKRTKIRKTKYTTILDNMDEQERMELRPKYGIKHRST 1020
 Db 961 VLAFTLIVLTGFTWLTCICCCCKRQKRTKIRKTKYTTILDNMDEQERMELRPKYGIKHRST 1020
 QY 1021 EHNSSLVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFSYCSKDR 1072
 Db 1021 EHNSSLVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFSYCSKDR 1072
 RESULT 2
 ADR00600
 ID ADR00600 standard; protein; 1072 AA.
 XX AC ADR00600;
 XX AC ADR00600;
 DT 04-NOV-2004 (first entry)
 XX Human 254P1D6B v.5 protein SEQ ID NO:11.
 DE 254P1D6B; small interfering RNA; siRNA; immune response;
 KW 254P1D6B-related protein; cytosolic; gene therapy; cancer; human;
 KW 254P1D6B v.5; chromosome 6.
 XX Homo sapiens.
 XX WO2004067716-A2.
 XX 12-AUG-2004.
 XX 23-JAN-2004; 2004WO-US001965.
 PR 24-JAN-2003; 2003US-0442526P.
 XX (AGEN-) AGENSYS INC.
 PI Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
 PI Perez-Villar JJ, Paris M;
 DR WPI; 2004-580991/56.
 XX New 254P1D6B siRNA composition comprising a double stranded siRNA that corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B protein or corresponds to a subsequence of the ORF, useful for detecting and treating cancer.
 PS Example 1; SEQ ID NO 11; 345pp; English.
 CC The present invention describes a 254P1D6B small interfering RNA (siRNA) composition that comprises a double stranded siRNA that corresponds to the nucleic acid open reading frame (ORF) sequence which encodes the 254P1D6B protein, or corresponds to a subsequence of the ORF, where the double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides in length. Also described: (1) a composition that comprises, consists essentially of, or consists of a peptide of eight, nine, ten, or eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID NO.3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in any of the 42 lists of peptides, given in the specification, or a protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or identical to an entire amino acid sequence of PI; (2) a polynucleotide that encodes the protein; (3) a composition comprising a polynucleotide that is a full complement of the polynucleotide described above; (4) generating a mammalian immune response directed to the protein of PI; (5) detecting, in a sample, the presence of a 254P1D6B-related protein or a

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:24:05 ; Search time 73 Seconds
(without alignments)
5679.550 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 5580
Sequence: 1 MAPPTGVLSLLLVTTIAGC.....VSMNGSIRNGASFSYCSKDR 1072

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 854512

Minimum DB seq length: 0
Maximum DB seq length: 35

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	158	2.8	29	8	ADR00857 Human 254
2	152	2.7	29	8	ADR00851 Human 254
3	101	1.8	19	8	ADR00850 Human 254
4	99	1.8	19	8	ADR00856 Human 254
5	92	1.6	16	8	ADR00632 Human 254
6	90	1.6	17	8	ADR00855 Human 254
7	84	1.5	17	8	ADR00849 Human 254
8	83	1.5	19	4	ABG22066 Novel hum
9	79	1.4	24	8	ADR00854 Human 254
10	63	1.1	9	8	ADR00727 Human 254
11	61.5	1.1	34	4	AB85615 Monoclonal
12	61.5	1.1	34	5	AB85625 Cerebrosp
13	60	1.1	12	8	ADR00630 Human 254
14	59	1.1	9	8	ADR00726 Human 254
15	59	1.1	9	8	ADR00838 Human 254
16	57	1.0	9	8	ADR00641 Human 254
17	57	1.0	31	3	AB25609 Peptide e
18	57	1.0	31	6	ADA27081 Human nov
19	57	1.0	31	8	ABE86613 Integrin
20	57	1.0	34	8	ABO57279 Human gen
21	56	1.0	9	8	ADR00814 Human 254
22	56	1.0	9	8	ADR00728 Human 254
23	56	1.0	9	8	ADR00829 Human 254
24	55	1.0	9	8	ADR00639 Human 254
25	55	1.0	9	8	ADR00675 Human 254

26	55	1.0	9	8	ADR00828 Human 254
27	55	1.0	11	8	ADR00631 Human 254
28	55	1.0	33	3	AAB25603 Peptide e
29	55	1.0	33	6	ADA27075 Human nov
30	55	1.0	33	8	ABE86607 Integrin
31	54.5	1.0	23	7	ADC35432 FITC label
32	54.5	1.0	23	7	ADC35439 FAM label
33	54	1.0	9	8	ADR00694 Human 254
34	54	1.0	9	8	ADR00826 Human 254
35	54	1.0	9	8	ADR00725 Human 254
36	54	1.0	9	8	ADR00747 Human 254
37	54	1.0	9	8	ADR00810 Human 254
38	54	1.0	29	1	AAP81554 Human ins
39	54	1.0	29	8	ADI80581 Conjugate
40	54	1.0	34	3	AAB25607 Peptide e
41	54	1.0	34	6	ADA27079 Human nov
42	54	1.0	34	8	ABE86611 Integrin
43	53.5	1.0	24	5	ABB08283 Human Gal
44	53.5	1.0	24	6	ABU07534 Human N-a
45	53.5	1.0	24	7	ADE25869 Human muc

ALIGNMENTS

RESULT 1

ADR00857 ID ADR00857 standard; peptide; 29 AA.
XX
AC ADR00857;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human 254PID6B v.5 peptide SEQ ID NO:268.
XX
KW 254PID6B; small interfering RNA; siRNA; immune response;
KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
KW 254PID6B v.5; chromosome 6.
XX
OS Homo sapiens.
XX
PN W02004067716-A2.
XX
PD 12-AUG-2004.
XX
PF 23-JAN-2004; 2004WO-US001965.
XX
PR 24-JAN-2003; 2003US-0442526P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
XX
PI Perez-Villar JJ, Faris M;
XX
DR WPI; 2004-580991/56.
XX
PT New 254PID6B siRNA composition comprising a double stranded siRNA that corresponds to the nucleic acid ORF sequence which encodes the 254PID6B protein or corresponds to a subsequence of the ORF, useful for detecting and treating cancer.
XX
PS Example 13; SEQ ID NO 268; 345pp; English.
XX

The present invention describes a 254PID6B small interfering RNA (siRNA) composition that comprises a double stranded siRNA that corresponds to the nucleic acid open reading frame (ORF) sequence which encodes the 254PID6B protein, or corresponds to a subsequence of the ORF, where the double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides in length. Also described: (1) a composition that comprises, consists essentially of, or consists of a peptide of eight, nine, ten, or eleven contiguous amino acids of a protein of figure 2 (p1, see SEQ ID NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in any of the 42 lists of peptides, given in the specification, or a protein

CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents a
 CC human 254P1D6B v.5 peptide, which is used in the exemplification of the
 CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX Sequence 29 AA;

Query Match 2.8%; Score 158; DB 8; Length 29;
 Best Local Similarity 96.6%; Pred. No. 0.00021;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 128 GIWGSPEDIRKDLFLGKDWGLEMSY 156
 |||||
 DB 1 GIWGSPEDIRKDLFLGKDWGLEMSY 29

RESULT 2

ADNR00851
 ID ADNR00851 standard; peptide; 29 AA.

XX AC ADNR00851;

XX DT 04-NOV-2004 (first entry)

XX DE Human 254P1D6B v.2 peptide SEQ ID NO:262.

XX KW 254P1D6B; small interfering RNA; siRNA; immune response;

XX KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;

XX KW 254P1D6B v.2; chromosome 6.

XX OS Homo sapiens.

XX PN WO2004067716-A2.

XX PD 12-AUG-2004.

XX PF 23-JAN-2004; 2004WO-US001965.

XX PR 24-JAN-2003; 2003US-0442526P.

XX PA (AGEN-) AGENSYS INC.

XX PI Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;

XX PI Perez-Villar JJ, Faris M;

XX DR WPI; 2004-580991/56.

XX PT New 254P1D6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

XX PS Example 13; SEQ ID NO 262; 345pp; English.

XX CC The present invention describes a 254P1D6B small interfering RNA (siRNA)

CC composition that comprises a double stranded siRNA that corresponds to

CC the nucleic acid open reading frame (ORF) sequence which encodes the

CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the

CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
 CC NO:3, 5 or 7 ADNR00592, ADNR00594 or ADNR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents a
 CC human 254P1D6B v.2 peptide, which is used in the exemplification of the
 CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX Sequence 29 AA;

Query Match 2.7%; Score 152; DB 8; Length 29;
 Best Local Similarity 96.6%; Pred. No. 0.00059;
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 143 FLGKDWGLEMSYKDDYRELEKDLQPS 171
 |||||

DB 1 FLGKDWGLEMSYKDDYRELEKDLQPS 29

RESULT 3

ADNR00850

ID ADNR00850 standard; peptide; 19 AA.

XX AC ADNR00850;

XX DT 04-NOV-2004 (first entry)

XX DE Human 254P1D6B v.2 peptide SEQ ID NO:261.

XX KW 254P1D6B; small interfering RNA; siRNA; immune response;

XX KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;

XX KW 254P1D6B v.2; chromosome 6.

XX OS Homo sapiens.

XX PN WO2004067716-A2.

XX PD 12-AUG-2004.

XX PF 23-JAN-2004; 2004WO-US001965.

XX PR 24-JAN-2003; 2003US-0442526P.

XX PA (AGEN-) AGENSYS INC.

XX PI Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;

XX PI Perez-Villar JJ, Faris M;

XX DR WPI; 2004-580991/56.

XX PT New 254P1D6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

XX XX

PS Example 13; SEQ ID NO 261; 345pp; English.

XX The present invention describes a 254p1d6B small interfering RNA (siRNA) that corresponds to the nucleic acid ORF sequence which encodes the 254p1d6B protein or corresponds to a subsequence of the ORF, useful for detecting and treating cancer.

CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides in length. Also described: (1) a composition that comprises, consists essentially of, or consists of a peptide of eight, nine, ten, or eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in any of the 42 lists of peptides, given in the specification, or a protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or identical to an entire amino acid sequence of P1; (2) a polynucleotide that encodes the protein; (3) a composition comprising a polynucleotide that is a full complement of the polynucleotide described above; (4) generating a mammalian immune response directed to the protein of P1; (5) detecting, in a sample, the presence of a 254p1d6B-related protein or a 254p1d6B-related polynucleotide; (6) a composition that modulates the status of a cell that expresses a protein of P1; (7) a pharmaceutical composition that comprises the composition described above in a human unit dose form; (8) an antibody or its fragment, which is monoclonal; (9) a non-human transgenic animal that produces the antibody; (10) a hybridoma that produces the antibody; (11) delivering a cytotoxic agent or a diagnostic agent to a cell that expresses the protein of P1; and (12) inhibiting growth, reproduction or survival of cancer cells that express the protein of P1. 254p1d6B has cytostatic activity, and can be used in gene therapy. The compositions, molecules and methods are useful for treating and detecting cancer. The present sequence represents a human 254p1d6B v.2 peptide, which is used in the exemplification of the present invention. The human 254p1d6B gene is located on chromosome 6p22.

XX Sequence 19 AA;

Qy 148 WGLEEMSEYDDYRELEKD 166

Db 1 WGLEEMSEYDDYRELEKD 19

RESULT 4

ADRO0856

ID ADR00856 standard; peptide; 19 AA.

XX AC ADR00856;

XX 04-NOV-2004 (first entry)

XX Human 254p1d6B v.5 peptide SEQ ID NO:267.

XX 254p1d6B; small interfering RNA; siRNA; immune response;

KW 254p1d6B-related protein; cytostatic; gene therapy; cancer; human;

KW 254p1d6B v.5; chromosome 6.

XX OS Homo sapiens.

XX WO2004067716-A2.

XX 12-AUG-2004.

XX 23-JAN-2004; 2004WO-US001965.

XX 24-JAN-2003; 2003US-0442526P.

XX (AGEN-) AGENSYS INC.

XX Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;

PI Perez-Villar JJ, Paris M;

XX WPI; 2004-580991/56.

XX Example 13; SEQ ID NO 267; 345pp; English.

XX The present invention describes a 254p1d6B small interfering RNA (siRNA) that corresponds to the nucleic acid ORF sequence which encodes the 254p1d6B protein or corresponds to a subsequence of the ORF, useful for detecting and treating cancer.

CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides in length. Also described: (1) a composition that comprises, consists essentially of, or consists of a peptide of eight, nine, ten, or eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in any of the 42 lists of peptides, given in the specification, or a protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or identical to an entire amino acid sequence of P1; (2) a polynucleotide that encodes the protein; (3) a composition comprising a polynucleotide that is a full complement of the polynucleotide described above; (4) generating a mammalian immune response directed to the protein of P1; (5) detecting, in a sample, the presence of a 254p1d6B-related protein or a 254p1d6B-related polynucleotide; (6) a composition that modulates the status of a cell that expresses a protein of P1; (7) a pharmaceutical composition that comprises the composition described above in a human unit dose form; (8) an antibody or its fragment, which is monoclonal; (9) a non-human transgenic animal that produces the antibody; (10) a hybridoma that produces the antibody; (11) delivering a cytotoxic agent or a diagnostic agent to a cell that expresses the protein of P1; and (12) inhibiting growth, reproduction or survival of cancer cells that express the protein of P1. 254p1d6B has cytostatic activity, and can be used in gene therapy. The compositions, molecules and methods are useful for treating and detecting cancer. The present sequence represents a human 254p1d6B v.5 peptide, which is used in the exemplification of the present invention. The human 254p1d6B gene is located on chromosome 6p22.

XX Sequence 19 AA;

Qy 133 SPEDIRKDLXFLGKDWGLE 151

Db 1 SPEDIRKDLXFLGKDWGLE 19

RESULT 5

ADRO0632

ID ADR00632 standard; peptide; 16 AA.

XX AC ADR00632;

XX 04-NOV-2004 (first entry)

XX Human 254p1d6B peptide SEQ ID NO:43.

XX 254p1d6B; small interfering RNA; siRNA; immune response;

KW 254p1d6B-related protein; cytostatic; gene therapy; cancer; human;

KW chromosome 6.

XX OS Homo sapiens.

XX WO2004067716-A2.

XX 12-AUG-2004.

XX 23-JAN-2004; 2004WO-US001965.

XX 24-JAN-2003; 2003US-0442526P.

XX Query Match 1.8%; Score 99; DB 8; Length 19;

Best Local Similarity 94.7%; Pred. No. 2.3;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
PA (AGEN-) AGENSYS INC.
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
PI Perez-Villar JJ, Paris M;
XX WPI; 2004-580991/56.
DR
XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
PS
XX Example 9; SEQ ID NO 43; 345pp; English.
XX
XX The present invention describes a 254P1D6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents the
CC human 254P1D6B peptide, which is used in the exemplification of the
CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
XX
SQ Sequence 16 AA;
Query Match 1.6%; Score 92; DB 8; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 522 NNAVDPYPVANAGPNH 537
Db 1 NNAVDPYPVANAGPNH 16
RESULT 6
ADR00855
ID ADR00855 standard; peptide; 17 AA.
XX
AC ADR00855;
XX
XX 04-NOV-2004 (first entry)
XX
XX Human 254P1D6B v.5 peptide SEQ ID NO:266.
XX
XX 254P1D6B; small interfering RNA; siRNA; immune response;
KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
KW 254P1D6B v.5; chromosome 6.
XX
XX Homo sapiens.
OS
XX WO2004067716-A2.
XX
XX
```

```
PD 12-AUG-2004.
XX
XX 23-JAN-2004; 2004WO-US001965.
XX
XX 24-JAN-2003; 2003US-0442526P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
PI Perez-Villar JJ, Paris M;
XX WPI; 2004-580991/56.
DR
XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
PS
XX Example 13; SEQ ID NO 266; 345pp; English.
XX
XX The present invention describes a 254P1D6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents a
CC human 254P1D6B v.5 peptide, which is used in the exemplification of the
CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
XX
SQ Sequence 17 AA;
Query Match 1.6%; Score 90; DB 8; Length 17;
Best Local Similarity 94.1%; Pred. No. 8.7;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 134 PEDIRKDLXFLGKDWGL 150
Db 1 PEDIRKDLTFLGKDWGL 17
RESULT 7
ADR00849
ID ADR00849 standard; peptide; 17 AA.
XX
AC ADR00849;
XX
XX 04-NOV-2004 (first entry)
XX
XX Human 254P1D6B v.2 peptide SEQ ID NO:260.
XX
XX 254P1D6B; small interfering RNA; siRNA; immune response;
KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
KW
```

KW 254P1D6B v.2; chromosome 6.
 XX Homo sapiens.
 OS
 XX WO2004067716-A2.
 FN
 XX 12-AUG-2004.
 PD
 XX
 XX 23-JAN-2004; 2004WO-US001965.
 PF
 XX 24-JAN-2003; 2003US-0442526P.
 PR
 XX (AGEN-) AGENSYS INC.
 PA
 XX Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
 FI Perez-Villar JJ, Paris M;
 PI
 XX WPI; 2004-580991/56.
 DR
 XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
 FT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.
 XX
 PS Example 13; SEQ ID NO 260; 345pp; English.
 XX
 CC The present invention describes a 254P1D6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents a
 CC human 254P1D6B v.2 peptide, which is used in the exemplification of the
 CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX
 SQ Sequence 17 AA;
 Query Match 1.5%; Score 84; DB 8; Length 17;
 Best Local Similarity 94.1%; Pred. No. 24;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 149 GLEEMSEYDDYRELEK 165
 Db 1 GLEEMSEYADYRELEK 17
 |||||
 RESULT 8
 ABG22066
 ID ABG22066 standard; protein; 19 AA.
 XX
 AC ABG22066;
 XX
 DT
 XX 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #22057.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS
 XX Homo sapiens.
 XX WO200175067-A2.
 FN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 XX 23-AUG-2000; 2000US-00649167.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 FI
 XX WPI; 2001-639362/73.
 DR
 XX N-PSDB; AAS86253.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 FT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 52425; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 19 AA;
 Query Match 1.5%; Score 83; DB 4; Length 19;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MAPPTGVLSLLLVTTIA 18
 Db 1 MAPPTGVLSLLLVTTIA 18
 |||||
 RESULT 9
 ADR00854
 ID ADR00854 standard; peptide; 24 AA.
 XX
 AC ADR00854;
 XX
 DT 04-NOV-2004 (first entry)
 XX

DE Human 254PID6B v.3 peptide SEQ ID NO:265.
XX 254PID6B; small interfering RNA; siRNA; immune response;
KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
XX 254PID6B v.3; chromosome 6.
OS Homo sapiens.
XX
XX
PN WO2004067716-A2.
XX 12-AUG-2004.
XX
XX 23-JAN-2004; 2004WO-US001965.
XX 24-JAN-2003; 2003US-0442526P.
XX (AGEN-) AGENSYS INC.
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
PI Perez-Villar JJ, Faris M;
XX WPI; 2004-580991/56.
XX
XX New 254PID6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX
XX Example 13; SEQ ID NO 265; 345pp; English.
XX
XX The present invention describes a 254PID6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (p1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of p1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of p1; (5)
CC detecting, in a sample, the presence of a 254PID6B-related protein or a
CC 254PID6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of p1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of p1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of p1. 254PID6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents a
CC human 254PID6B v.3 peptide, which is used in the exemplification of the
CC present invention. The human 254PID6B gene is located on chromosome 6p22.
XX
XX Sequence 24 AA;
Query Match 1.4%; Score 79; DB 8; Length 24;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 CARQKQCEGRTYSN 33
Db 11 CARQKQCEGRTYSN 24
|||||
RESULT 10
ADR00727

ADR00727 standard; peptide; 9 AA.
XX
XX ADR00727;
XX
DT 04-NOV-2004 (first entry)
XX
XX Human 254PID6B peptide SEQ ID NO:138.
XX
XX 254PID6B; small interfering RNA; siRNA; immune response;
KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
XX chromosome 6.
XX
XX Homo sapiens.
XX OS
XX WO2004067716-A2.
XX
XX 12-AUG-2004.
XX
XX 23-JAN-2004; 2004WO-US001965.
XX 24-JAN-2003; 2003US-0442526P.
XX (AGEN-) AGENSYS INC.
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
PI Perez-Villar JJ, Faris M;
XX WPI; 2004-580991/56.
XX
XX New 254PID6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX
XX Example 9; SEQ ID NO 138; 345pp; English.
XX
XX The present invention describes a 254PID6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (p1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of p1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of p1; (5)
CC detecting, in a sample, the presence of a 254PID6B-related protein or a
CC 254PID6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of p1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of p1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of p1. 254PID6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents the
CC human 254PID6B peptide, which is used in the exemplification of the
CC present invention. The human 254PID6B gene is located on chromosome 6p22.
XX
XX Sequence 9 AA;
Query Match 1.1%; Score 63; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 946 WDGSNCWEW 954

SQ Sequence 34 AA;
 Query Match 1.1%; Score 61.5; DB 5; Length 34;
 Best Local Similarity 29.6%; Pred. No. 3.2e+03;
 Matches 21; Conservative 6; Mismatches 7; Indels 37; Gaps 4;
 QY 176 PRGSAEYTDWGLLPGEAGFNSVGDSPAVPAETQDDPELHYLNSASTPAPKLPERSVL 235
 DB 1 PRGAAP-----PQKGQANAT-----RIPAKT-----PPAPKTB----- 29
 QY 236 LPLPTTPSSGE 246
 DB 30 -----PSSGE 34
 RESULT 13
 ADR00630
 ID ADR00630 standard; peptide; 12 AA.
 AC ADR00630;
 DT 04-NOV-2004 (first entry)
 DE Human 254P1D6B peptide SEQ ID NO:41.
 XX
 KW 254P1D6B; small interfering RNA; siRNA; immune response;
 KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
 KW chromosome 6.
 XX Homo sapiens.
 OS
 XX WO2004067716-A2.
 XX 12-AUG-2004.
 XX
 PF 23-JAN-2004; 2004WO-US001965.
 XX
 PR 24-JAN-2003; 2003US-0442526P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Faris M;
 XX
 DR WPI; 2004-580991/56.
 XX
 PT New 254P1D6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.
 XX
 PS Example 9; SEQ ID NO 41; 345pp; English.
 XX
 CC The present invention describes a 254P1D6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of PI; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of PI; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of PI; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)

a non-human transgenic animal that produces the antibody; (10) a
 hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 or a diagnostic agent to a cell that expresses the protein of PI; and
 (12) inhibiting growth, reproduction or survival of cancer cells that
 express the protein of PI. 254P1D6B has cytostatic activity, and can be
 used in gene therapy. The compositions, molecules and methods are useful
 for treating and detecting cancer. The present sequence represents the
 human 254P1D6B peptide, which is used in the exemplification of the
 present invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX Sequence 12 AA;
 Query Match 1.1%; Score 60; DB 8; Length 12;
 Best Local Similarity 91.7%; Pred. No. 7.8e+02;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 151 EEMSEYXDDYRE 162
 DB 1 EEMSEYSDDYRE 12
 RESULT 14
 ADR00726
 ID ADR00726 standard; peptide; 9 AA.
 XX
 AC ADR00726;
 DT 04-NOV-2004 (first entry)
 DE Human 254P1D6B peptide SEQ ID NO:137.
 XX
 KW 254P1D6B; small interfering RNA; siRNA; immune response;
 KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
 KW chromosome 6.
 XX Homo sapiens.
 OS
 XX WO2004067716-A2.
 XX 12-AUG-2004.
 XX
 PF 23-JAN-2004; 2004WO-US001965.
 XX
 PR 24-JAN-2003; 2003US-0442526P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Faris M;
 XX
 DR WPI; 2004-580991/56.
 XX
 PT New 254P1D6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.
 XX
 PS Example 9; SEQ ID NO 137; 345pp; English.
 XX
 CC The present invention describes a 254P1D6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of PI; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)

CC Generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254P1D6B peptide, which is used in the exemplification of the
 CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX
 SQ Sequence 9 AA;
 Query Match 1.1%; Score 59; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 929 RCICSHLWM 937
 DB 1 RCICSHLWM 9
 RESULT 15
 ADNR0838
 ID ADR00838 standard; peptide; 9 AA.
 XX
 AC ADR00838;
 XX
 XX 04-NOV-2004 (first entry)
 DT
 XX Human 254P1D6B peptide SEQ ID NO:249.
 DE
 XX 254P1D6B; small interfering RNA; siRNA; immune response;
 KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
 KW chromosome 6.
 KW
 XX Homo sapiens.
 OS
 XX WO2004067716-A2.
 PN
 XX 12-AUG-2004.
 PD
 XX 23-JAN-2004; 2004WO-US001965.
 PF
 XX 24-JAN-2003; 2003US-0442526P.
 PR
 XX (AGEN-) AGENSYS INC.
 PA
 XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Paris M;
 XX
 XX WPI; 2004-580991/56.
 DR
 XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.
 PT
 XX Example 9; SEQ ID NO 249; 345pp; English.
 PS
 XX The present invention describes a 254P1D6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID

CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254P1D6B peptide, which is used in the exemplification of the
 CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX
 SQ Sequence 9 AA;
 Query Match 1.1%; Score 59; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 654 GIVFYHWEH 662
 DB 1 GIVFYHWEH 9
 Search completed: October 12, 2005, 10:29:17
 Job time : 75 secs

254P1D6B-related polynucleotide; (6) a composition that modulates the status of a cell that expresses a protein of P1; (7) a pharmaceutical composition that comprises the composition described above in a human unit dose form; (8) an antibody or its fragment, which is monoclonal; (9) a non-human transgenic animal that produces the antibody; (10) a hybridoma that produces the antibody; (11) delivering a cytotoxic agent or a diagnostic agent to a cell that expresses the protein of P1; and (12) inhibiting growth, reproduction or survival of cancer cells that express the protein of P1. 254P1D6B has cytostatic activity, and can be used in gene therapy. The compositions, molecules and methods are useful for treating and detecting cancer. The present sequence represents the human 254P1D6B v.5, which is used in the exemplification of the present invention. The human 254P1D6B gene is located on chromosome 6p22.

XX Sequence 1072 AA;

Query Match 99.9%; Score 5576; DB 8; Length 1072;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAPPTGVLSSLLLVTTIAGCARKQCSGRYTSNAVISPNLETTIRMRVSHTPPVVDCTAA 60
 Db 1 MAPPTGVLSSLLLVTTIAGCARKQCSGRYTSNAVISPNLETTIRMRVSHTPPVVDCTAA 60

Qy 61 CCDSLLDAAWFGRCYLVSPCHKENCEPKMGPIRSYLTFLVRPVPORPAQLLDYGDMM 120
 Db 61 CCDSLLDAAWFGRCYLVSPCHKENCEPKMGPIRSYLTFLVRPVPORPAQLLDYGDMM 120

Qy 121 LNRGSPGIMGDSPEIRKDLKFLGKDWGLMESEYDDYRELEKDLQPSGKQEPGRSA 180
 Db 121 LNRGSPGIMGDSPEIRKDLKFLGKDWGLMESEYDDYRELEKDLQPSGKQEPGRSA 180

Qy 181 EYTWGLLPGEAFNSVGSVPAPVAPQDQDELHYLNESASTPAPKLPERSVLLPLPT 240
 Db 181 EYTWGLLPGEAFNSVGSVPAPVAPQDQDELHYLNESASTPAPKLPERSVLLPLPT 240

Qy 241 TPSSGEVLEKAKAQLOEQSSNSGKVEVLPMSHSLPPASLSLSSVTVKEKSPVLTWTGCT 300
 Db 241 TPSSGEVLEKAKAQLOEQSSNSGKVEVLPMSHSLPPASLSLSSVTVKEKSPVLTWTGCT 300

Qy 301 EHSIPTPTSAAPSESTPSELPIPTAPRTVKELTVSAGNLIITLPDNEVELKAFVAP 360
 Db 301 EHSIPTPTSAAPSESTPSELPIPTAPRTVKELTVSAGNLIITLPDNEVELKAFVAP 360

Qy 361 APPVETTYNENWLISHPTDYQGIKQKHQKTLNLSQLSVGLYFKVTVSSENAFGEFV 420
 Db 361 APPVETTYNENWLISHPTDYQGIKQKHQKTLNLSQLSVGLYFKVTVSSENAFGEFV 420

Qy 421 NVTVKPARRVNLPPVAVVSPQLBELTLPLTSALIDGQSSTDDTETVSYHWEIINGPFTIE 480
 Db 421 NVTVKPARRVNLPPVAVVSPQLBELTLPLTSALIDGQSSTDDTETVSYHWEIINGPFTIE 480

Qy 481 KTSVDSPLRLSNLDPGNSYFRLTWTDSGATNSTTAALIYNNAVDYPPVANAGNHTIT 540
 Db 481 KTSVDSPLRLSNLDPGNSYFRLTWTDSGATNSTTAALIYNNAVDYPPVANAGNHTIT 540

Qy 541 LPQNSITLNGQSDHDIQVLYEWSLPGSGEGKHVVMQGVQTPYLHLSAMQEGDYTFQLK 600
 Db 541 LPQNSITLNGQSDHDIQVLYEWSLPGSGEGKHVVMQGVQTPYLHLSAMQEGDYTFQLK 600

Qy 601 VTDSRQOSTAXVTIVOPENNRPPVAVAGDPKELIPVESATLDGSSSDHGHVFFVHW 660
 Db 601 VTDSRQOSTAXVTIVOPENNRPPVAVAGDPKELIPVESATLDGSSSDHGHVFFVHW 660

Qy 661 EHVGRPSAVENIDKATATVGLQVGYTHFRLTVKQOGLSSSTLTVAVKENNSPPR 720
 Db 661 EHVGRPSAVENIDKATATVGLQVGYTHFRLTVKQOGLSSSTLTVAVKENNSPPR 720

Qy 721 ARAGGRHVLLPNNSITLDGSRSTDDQIRIVSYLWIRDCQSPAAGDVIDGSDHVALQLTN 780
 Db 721 ARAGGRHVLLPNNSITLDGSRSTDDQIRIVSYLWIRDCQSPAAGDVIDGSDHVALQLTN 780

Qy 781 LVEGVYTHLVRVDSQGASDTRATVEVQDPDRKSGLVLTQLQVGQLTQRKDTLVRQ 840

Db 781 LVEGVYTHLVRVDSQGASDTRATVEVQDPDRKSGLVLTQLQVGQLTQRKDTLVRQ 840
 Qy 841 LAVLLNVLDSIDIKVKIRAHSDLSITVIVYVQSRPPFKVLKAAEVARNLHMLSKKADF 900
 Db 841 LAVLLNVLDSIDIKVKIRAHSDLSITVIVYVQSRPPFKVLKAAEVARNLHMLSKKADF 900

Qy 901 LLPKVLVDTAGCLLKCSGHCHCDPLTKRCICSHLWENLIQRYWIDGESNCSESIFFVT 960
 Db 901 LLPKVLVDTAGCLLKCSGHCHCDPLTKRCICSHLWENLIQRYWIDGESNCSESIFFVT 960

Qy 961 VLAFTLVLGTGFTWLCICCCCKRQKTKIRKTKYTTILDNNDEQERMELRPKYGIKHRST 1020
 Db 961 VLAFTLVLGTGFTWLCICCCCKRQKTKIRKTKYTTILDNNDEQERMELRPKYGIKHRST 1020

Qy 1021 EHNSLMVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSYCKDR 1072
 Db 1021 EHNSLMVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSYCKDR 1072

RESULT 3
 ADR00601
 ID ADR00601 standard; protein; 1072 AA.
 XX ADR00601;
 AC ADR00601;
 DT 04-NOV-2004 (first entry)
 XX Human 254P1D6B v.6 protein SEQ ID NO:12.
 DE 254P1D6B; small interfering RNA; siRNA; immune response;
 KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
 KW 254P1D6B v.6; chromosome 6.
 XX Homo sapiens.
 XX WO2004067716-A2.
 XX 12-AUG-2004.
 XX 23-JAN-2004; 2004WO-US001965.
 XX 24-JAN-2003; 2003US-0442526P.
 XX (AGEN-) AGENSYS INC.
 XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Paris M;
 XX WPI; 2004-580991/56.
 DR New 254P1D6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
 FT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.
 XX Example 1; SEQ ID NO 12; 345pp; English.
 PS The present invention describes a 254P1D6B small interfering RNA (siRNA)
 XX composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)

CC detecting, in a sample, the presence of a 254p1d6B-related protein or a
 CC 254p1d6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254p1d6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254p1d6B v.6, which is used in the exemplification of the present
 CC invention. The human 254p1d6B gene is located on chromosome 6p22.
 XX
 SQ Sequence 1072 AA;

Query Match 99.9%; Score 5574; DB 8; Length 1072;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPPTGVLSLLLLVITAGCARKQCSGRTYSNAVISPNLETRIMRVSHTPPVVDTAA 60
 DB 1 MAPPTGVLSLLLLVITAGCARKQCSGRTYSNAVISPNLETRIMRVSHTPPVVDTAA 60
 QY 61 CDDLSCLAWFEGRCVLSVCPHKENCEPKMGPIRSVLTFLVPVORPAQLLDYGDMM 120
 DB 61 CDDLSCLAWFEGRCVLSVCPHKENCEPKMGPIRSVLTFLVPVORPAQLLDYGDMM 120
 QY 121 LNRGSPGSGIPEDIRKDLXFLGKMGLEMSYXDYRELEKDLQPSGQKPRGSA 180
 DB 121 LNRGSPGSGIPEDIRKDLXFLGKMGLEMSYVADDYRELEKDLQPSGQKPRGSA 180
 QY 181 EYTDGGLLPGSGAFNNSVGDSPVAPETQDPPELHYNESASTAPKLPERSVLLPPT 240
 DB 181 EYTDGGLLPGSGAFNNSVGDSPVAPETQDPPELHYNESASTAPKLPERSVLLPPT 240
 QY 241 TPSSGEVLEKAKASQLOEQSSNSGKEVLMPSHSLPPASLELSSVTVKESPVLTTPGST 300
 DB 241 TPSSGEVLEKAKASQLOEQSSNSGKEVLMPSHSLPPASLELSSVTVKESPVLTTPGST 300
 QY 301 EHSIPTPTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
 DB 301 EHSIPTPTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
 QY 361 APPEVTTNYEWNLSHPTDYQCEIKQGHKOTLNLSQLSVGLYVFKVTVSSNAFGEFV 420
 DB 361 APPEVTTNYEWNLSHPTDYQCEIKQGHKOTLNLSQLSVGLYVFKVTVSSNAFGEFV 420
 QY 421 NVTVPARVNLPPVAVVSPQQLTLPLTSALIDGSQSTDDTEIVSVHWEINGPFTIE 480
 DB 421 NVTVPARVNLPPVAVVSPQQLTLPLTSALIDGSQSTDDTEIVSVHWEINGPFTIE 480
 QY 481 KTSVDSVPLRLSNLDPGNYSFRLTVTSDGATNSTTAALIVNNAVDPYPVANAGNHTIT 540
 DB 481 KTSVDSVPLRLSNLDPGNYSFRLTVTSDGATNSTTAALIVNNAVDPYPVANAGNHTIT 540
 QY 541 LPQNSITLNGQSSDDHQIVLWELSLGPGSEKGVVMQGVQTPYHLNSAQSGDYTFQIK 600
 DB 541 LPQNSITLNGQSSDDHQIVLWELSLGPGSEKGVVMQGVQTPYHLNSAQSGDYTFQIK 600
 QY 601 VTDSSRQOSTAVTVTVQPNRRPPVAVAGPKELIFPVESATLDGSSSSDDHGIVFVHW 660
 DB 601 VTDSSRQOSTAVTVTVQPNRRPPVAVAGPKELIFPVESATLDGSSSSDDHGIVFVHW 660
 QY 661 EHVGPASAVEMENIDKAIATVGLQVGYHFLTVKDKQGLSSTSTLTAVVKKENNSPPR 720
 DB 661 EHVGPASAVEMENIDKAIATVGLQVGYHFLTVKDKQGLSSTSTLTAVVKKENNSPPR 720
 QY 721 ARAGGRHVLVLPNNSITLDGSRSTDQRIVSVLWIRDGQSPAGDVIDGSDHSVALQLTN 780
 DB 721 ARAGGRHVLVLPNNSITLDGSRSTDQRIVSVLWIRDGQSPAGDVIDGSDHSVALQLTN 780

QY 781 LVEGYTFHLRVTDSDQASDSTDTATVEVQDPDRKSLGVELTLOVGVGQLTQORKDTLVRQ 840
 DB 781 LVEGYTFHLRVTDSDQASDSTDTATVEVQDPDRKSLGVELTLOVGVGQLTQORKDTLVRQ 840
 QY 841 LAVLLNVLDSDIKVQKIRAHSDLSSTVIVFYVQSPPPFKVLAEEVARNLHRLSKEKADF 900
 DB 841 LAVLLNVLDSDIKVQKIRAHSDLSSTVIVFYVQSPPPFKVLAEEVARNLHRLSKEKADF 900
 QY 901 LLEFVLRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCEWSIFVYT 960
 DB 901 LLEFVLRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCEWSIFVYT 960
 QY 961 VLAFTLIVLTGCGFTWLCICCCCKRQKRTKIRKTKYTTILDNDMDQERBELPKYGIKHRST 1020
 DB 961 VLAFTLIVLTGCGFTWLCICCCCKRQKRTKIRKTKYTTILDNDMDQERBELPKYGIKHRST 1020
 QY 1021 EHNSSLMVSESEFSDSDQDTIFSRKMERGNPKVSMNGSIRNGASFSYCSKDR 1072
 DB 1021 EHNSSLMVSESEFSDSDQDTIFSRKMERGNPKVSMNGSIRNGASFSYCSKDR 1072
 RESULT 4
 ADR00597
 ID ADR00597 standard; protein; 1072 AA.
 XX
 AC ADR00597;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Human 254p1d6B v.1 clone LCP-3 protein SEQ ID NO:8.
 XX
 KW 254p1d6B; small interfering RNA; siRNA; immune response;
 KW 254p1d6B-related protein; cytostatic; gene therapy; cancer; human;
 KW 254p1d6B v.1 clone LCP-3; chromosome 6.
 XX
 OS Homo sapiens.
 XX
 PN WO2004067716-A2.
 XX
 PD 12-AUG-2004.
 XX
 PF 23-JAN-2004; 2004WO-US001965.
 XX
 PR 24-JAN-2003; 2003US-0442526P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Paris M;
 XX
 WPI; 2004-580991/56.
 XX
 PT New 254p1d6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254p1d6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.
 XX
 PS Example 1; SEQ ID NO 8; 345pp; English.
 XX
 CC The present invention describes a 254p1d6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254p1d6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)

CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents the
CC human 254P1D6B v.1 clone LCP-3, which is used in the exemplification of
CC the present invention. The human 254P1D6B gene is located on chromosome
CC 6p22.
XX
SQ Sequence 1072 AA;
Query Match 99.9%; Score 5574; DB 8; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MAPPTGVLSLLLLVTTIAGCARKQCSEGRYSNAVISPNLETTIRIMRVSHTFPPVVDCTAA 60
Db 1 MAPPTGVLSLLLLVTTIAGCARKQCSEGRYSNAVISPNLETTIRIMRVSHTFPPVVDCTAA 60
Qy 61 CCCLSSCDLAWFPGRCYLVSCPHKENCEPKMGPIRSYLTFLVRPQVPAQLLDYDGM 120
Db 61 CCCLSSCDLAWFPGRCYLVSCPHKENCEPKMGPIRSYLTFLVRPQVPAQLLDYDGM 120
Qy 121 LNRGSPGSGWDSPEIRKDLXFLGKOWGLEMEFYDDYRELEKDLLQPSGKQPRGSA 180
Db 121 LNRGSPGSGWDSPEIRKDLXFLGKOWGLEMEFYDDYRELEKDLLQPSGKQPRGSA 180
Qy 181 EYTDWGLLPGEAGFNSVGVSPVAPVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
Db 181 EYTDWGLLPGEAGFNSVGVSPVAPVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
Qy 241 TPSSGEVLEKEKASQLQEQSSNSGKEYLMPHSLPASPASLESSTVVEKSPVLTVTGST 300
Db 241 TPSSGEVLEKEKASQLQEQSSNSGKEYLMPHSLPASPASLESSTVVEKSPVLTVTGST 300
Qy 301 EHSITPTPTSAPESTPSELPIPTTAPRTVKELTVSAGNLIITLPDNEVELKAFVAP 360
Db 301 EHSITPTPTSAPESTPSELPIPTTAPRTVKELTVSAGNLIITLPDNEVELKAFVAP 360
Qy 361 APPVETTYEWNLIHSHTDYQGEIKQGHKOTLNLSQLSVGLYVFKVTVSSENAPGEGFV 420
Db 361 APPVETTYEWNLIHSHTDYQGEIKQGHKOTLNLSQLSVGLYVFKVTVSSENAPGEGFV 420
Qy 421 NVTVPARRVNLPPVAVVSPQLQBLTLPPLTSALIDGSGSTDDTEIVSYHWEENGPTEE 480
Db 421 NVTVPARRVNLPPVAVVSPQLQBLTLPPLTSALIDGSGSTDDTEIVSYHWEENGPTEE 480
Qy 481 KTSVDSVPLRSLNLDPGNYSFRLTVTDSGATNSTTAALIIVNNAVDPVPVANAGNHTIT 540
Db 481 KTSVDSVPLRSLNLDPGNYSFRLTVTDSGATNSTTAALIIVNNAVDPVPVANAGNHTIT 540
Qy 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGKGVVMQGVQTPVYLHLSAMQEGDYTFQLK 600
Db 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGKGVVMQGVQTPVYLHLSAMQEGDYTFQLK 600
Qy 601 VTDSRQOSTAXVTIVOPENNRPPVAVAGDPDKELIPFVESATLDGSSSSDDHGIVFVHW 660
Db 601 VTDSRQOSTAXVTIVOPENNRPPVAVAGDPDKELIPFVESATLDGSSSSDDHGIVFVHW 660
Qy 661 EHVGPASAVENIDKAIATVTLQVGYTHFRLTVKDDQGLSSTLTVAVKKENNSPPR 720
Db 661 EHVGPASAVENIDKAIATVTLQVGYTHFRLTVKDDQGLSSTLTVAVKKENNSPPR 720
Qy 721 ARAGGRHVLPNNISITLDGSRSTDQRIYSLWIRDOQSPAAGDVIDGSDHVALQLTN 780
Db 721 ARAGGRHVLPNNISITLDGSRSTDQRIYSLWIRDOQSPAAGDVIDGSDHVALQLTN 780

Db 721 ARAGGRHVLPNNISITLDGSRSTDQRIYSLWIRDOQSPAAGDVIDGSDHVALQLTN 780
Qy 781 LVEGVYTHLRVTDQSASDTDTATVENVQDPDRKSLGVLVELTLQVGVGQLTQRKDTLVRQ 840
Db 781 LVEGVYTHLRVTDQSASDTDTATVENVQDPDRKSLGVLVELTLQVGVGQLTQRKDTLVRQ 840
Qy 841 LAVLLNVLDSDIKVKIRAHSDLSLTVIVFYVQSRPPPKVLKAAEVARNLHMLRSLKEKADF 900
Db 841 LAVLLNVLDSDIKVKIRAHSDLSLTVIVFYVQSRPPPKVLKAAEVARNLHMLRSLKEKADF 900
Qy 901 LLFKVLRVDTAGCLLKCSGHGHCPLTKRCICSHLWMENLIQRYIWDGESNCEWSIFVYT 960
Db 901 LLFKVLRVDTAGCLLKCSGHGHCPLTKRCICSHLWMENLIQRYIWDGESNCEWSIFVYT 960
Qy 961 VLAFTLVLVGGFTWLCICCCCKRQRTKIRKTKYTTILDNDQERMELRPKYGIKHRST 1020
Db 961 VLAFTLVLVGGFTWLCICCCCKRQRTKIRKTKYTTILDNDQERMELRPKYGIKHRST 1020
Qy 1021 EHNSSLMVSESEPDSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSYCKDR 1072
Db 1021 EHNSSLMVSESEPDSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSYCKDR 1072
RESULT 5
ADRO0598
ID ADRO0598 standard; protein; 1072 AA.
XX ADRO0598;
XX AC ADRO0598;
XX DT 04-NOV-2004 (first entry)
XX DE Human 254P1D6B v.2 protein SEQ ID NO:9.
XX KW 254P1D6B; small interfering RNA; siRNA; immune response;
XX KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
XX KW 254P1D6B v.2; chromosome 6.
XX OS Homo sapiens.
XX WO2004067716-A2.
XX PD 12-AUG-2004.
XX PF 23-JAN-2004; 2004WO-US001965.
XX PR 24-JAN-2003; 2003US-0442526P.
XX PA (AGEN-) AGENSYS INC.
XX PI Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
PI Perez-Villar JJ, Paris M;
XX DR WPI; 2004-580991/56.
XX PT New 254P1D6B siRNA composition comprising a double stranded siRNA that
XX corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
XX protein or corresponds to a subsequence of the ORF, useful for detecting
XX and treating cancer.
XX Example 1; SEQ ID NO 9; 345pp; English.
XX The present invention describes a 254P1D6B small interfering RNA (siRNA)
XX composition that comprises a double stranded siRNA that corresponds to
XX the nucleic acid open reading frame (ORF) sequence which encodes the
XX 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
XX double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
XX nucleotides in length. Also described: (1) a composition that comprises,
XX consists essentially of, or consists of a peptide of eight, nine, ten, or
XX eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
XX NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
XX any of the 42 lists of peptides, given in the specification, or a protein
XX that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
XX identical to an entire amino acid sequence of P1; (2) a polynucleotide

CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents the
CC human 254P1D6B v.2, which is used in the exemplification of the present
CC invention. The human 254P1D6B gene is located on chromosome 6p22.
XX
SQ Sequence 1072 AA;
Query Match 99.9%; Score 5574; DB 8; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAPPTGVLSLLLVTTIAGCARKQCEGRTYGNVISPNIETTRIMRVSHTPPVVDCTAA 60
DB 1 MAPPTGVLSLLLVTTIAGCARKQCEGRTYGNVISPNIETTRIMRVSHTPPVVDCTAA 60
QY 61 CDDLSCLAWFEGRCYLVSCHKENCPEKMGPIRSYLTFLVLPVQRPQAQLLDYGDMM 120
DB 61 CDDLSCLAWFEGRCYLVSCHKENCPEKMGPIRSYLTFLVLPVQRPQAQLLDYGDMM 120
QY 121 LNRGSPSGTWGSDPEDIRKDLFLGKWGLEMSYXDYRELEKDLQPSKQKPRGSA 180
DB 121 LNRGSPSGTWGSDPEDIRKDLFLGKWGLEMSYADYRELEKDLQPSKQKPRGSA 180
QY 181 EYTDWGLLPGSGAFNSVGDSPVAPAEQOQDPHELHYNESASTAPKLPERSVLLPPT 240
DB 181 EYTDWGLLPGSGAFNSVGDSPVAPAEQOQDPHELHYNESASTAPKLPERSVLLPPT 240
QY 241 TPSSGEVLEKASQALQEQSSNSGKEVLMPSHSLPPASLELSVTVTEKSPVLVTPGST 300
DB 241 TPSSGEVLEKASQALQEQSSNSGKEVLMPSHSLPPASLELSVTVTEKSPVLVTPGST 300
QY 301 EHSIPTPTSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLITLPDNEVELKAFVAP 360
DB 301 EHSIPTPTSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLITLPDNEVELKAFVAP 360
QY 361 APPVETTYNENWNLISHPTDYOGEIKQGHKQTLNLSQLSVGLVYFVKTVSSNAFGEGRV 420
DB 361 APPVETTYNENWNLISHPTDYOGEIKQGHKQTLNLSQLSVGLVYFVKTVSSNAFGEGRV 420
QY 421 NVTVPARRVNLPPVAVVSQQLTLPITSALIDGQSQTDDTEIVSYHWEESINGPFIEE 480
DB 421 NVTVPARRVNLPPVAVVSQQLTLPITSALIDGQSQTDDTEIVSYHWEESINGPFIEE 480
QY 481 KTSVDSPLRLNLDPCGNSFRLLTVTDSGATNSTTAAIIVNNAVDPYPVANAGNHTTIT 540
DB 481 KTSVDSPLRLNLDPCGNSFRLLTVTDSGATNSTTAAIIVNNAVDPYPVANAGNHTTIT 540
QY 541 LPONSIITLNGNSSDDHQVLVLEWSLGPSEGHVVMQGVQVQPYLHLSAQMGSDYTFOLK 600
DB 541 LPONSIITLNGNSSDDHQVLVLEWSLGPSEGHVVMQGVQVQPYLHLSAQMGSDYTFOLK 600
QY 601 VTDSSRQSQSTAXVTVTVQPPENRPPVAVAGPKELIFPVESATLDGSSSSDDHGIVFYHW 660
DB 601 VTDSSRQSQSTAXVTVTVQPPENRPPVAVAGPKELIFPVESATLDGSSSSDDHGIVFYHW 660
QY 661 EHVGRPSAVEMENIDKAIATVGLQVGYTHFRITVKDQOGLSTSTLTVAVKKNNSPPR 720
DB 661 EHVGRPSAVEMENIDKAIATVGLQVGYTHFRITVKDQOGLSTSTLTVAVKKNNSPPR 720
QY 721 ARAGGRHVLVLPNNISITLDGSRSTDQRIIVSYLWIRDGQSPAAGDIVDGS DSHVALQLTN 780

DB 721 ARAGGRHVLVLPNNISITLDGSRSTDQRIIVSYLWIRDGQSPAAGDIVDGS DSHVALQLTN 780
QY 781 LVEGVYTFHURVTDQSGASDTDTATVEQDPKRSGLVELTLQVGVGQLTQQRKDTLVRQ 840
DB 781 LVEGVYTFHURVTDQSGASDTDTATVEQDPKRSGLVELTLQVGVGQLTQQRKDTLVRQ 840
QY 841 LAVILNVLDSDIKVKIRAHSDLSLTVIVFYVQSRPPPFKVLKAAEYARNLHRLSKEKADF 900
DB 841 LAVILNVLDSDIKVKIRAHSDLSLTVIVFYVQSRPPPFKVLKAAEYARNLHRLSKEKADF 900
QY 901 LLFKVLVRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCSEWSIFVYT 960
DB 901 LLFKVLVRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCSEWSIFVYT 960
QY 961 VLAFTLVLVTGFTWLCICCCCKROKRTKIRKTKYTTILDNDDEQERMLRPKYGIKHRST 1020
DB 961 VLAFTLVLVTGFTWLCICCCCKROKRTKIRKTKYTTILDNDDEQERMLRPKYGIKHRST 1020
QY 1021 EHNSLSMVSESEFSDSDQDTIFSRKMERGNPKVSNMGSIRNGASFSYCSKDR 1072
DB 1021 EHNSLSMVSESEFSDSDQDTIFSRKMERGNPKVSNMGSIRNGASFSYCSKDR 1072
RESULT 6
ADRO0848
ID ADR00848 standard; protein; 1072 AA.
XX ADR00848;
XX 04-NOV-2004 (first entry)
XX Human 254P1D6B v.1 amino acid sequence SEQ ID NO:259.
XX 254P1D6B; small interfering RNA; siRNA; immune response;
KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
KW 254P1D6B v.1; chromosome 6.
XX Homo sapiens.
XX WO2004067716-A2.
XX 12-AUG-2004.
XX 23-JAN-2004; 2004WO-US001965.
XX 24-JAN-2003; 2003US-0442526P.
XX (AGEN-) AGENSYS INC.
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PW, Ge W;
PI Perez-Villar JJ, Paris M;
XX WPI; 2004-580991/56.
XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX Example 13; SEQ ID NO 259; 345pp; English.
PS The present invention describes a 254P1D6B small interfering RNA (siRNA)
XX composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or

identical to an entire amino acid sequence of p1; (2) a polynucleotide that encodes the protein; (3) a composition comprising a polynucleotide that is a full complement of the polynucleotide described above; (4) generating a mammalian immune response directed to the protein of p1; (5) detecting, in a sample, the presence of a 254p1D6B-related protein or a 254p1D6B-related polynucleotide; (6) a composition that modulates the status of a cell that expresses a protein of p1; (7) a pharmaceutical composition that comprises the composition described above in a human unit dose form; (8) an antibody or its fragment, which is monoclonal; (9) a non-human transgenic animal that produces the antibody; (10) a hybridoma that produces the antibody; (11) delivering a cytotoxic agent or a diagnostic agent to a cell that expresses the protein of p1; and (12) inhibiting growth, reproduction or survival of cancer cells that express the protein of p1. 254p1D6B has cytostatic activity, and can be used in gene therapy. The compositions, molecules and methods are useful for treating and detecting cancer. The present sequence represents a human 254p1D6B v.1 amino acid sequence, which is used in the exemplification of the present invention. The human 254p1D6B gene is located on chromosome 6p22.

XX Seq Sequence 1072 AA;

Query Match 99.9%; Score 5574; DB 8; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAPPTGVLSLLLVTTIAGCARKQCSSEGRVTSNAVISPNLETTIRMRVSHVTPVVDCTAA 60
Db 1 MAPPTGVLSLLLVTTIAGCARKQCSSEGRVTSNAVISPNLETTIRMRVSHVTPVVDCTAA 60
Qy 61 CCDSLCDLAWFGRGCVLSVCPHKECEPKMGPIRSYLTFLVRPVRPAQLLDYDGM 120
Db 61 CCDSLCDLAWFGRGCVLSVCPHKECEPKMGPIRSYLTFLVRPVRPAQLLDYDGM 120
Qy 121 LNRGSPGIMGDSPEIRKDLKFLGKDWGLMESEYDDYRELEKDLQLQPSGKQEPGRSA 180
Db 121 LNRGSPGIMGDSPEIRKDLKFLGKDWGLMESEYDDYRELEKDLQLQPSGKQEPGRSA 180
Qy 181 EYTDWGLLPSEGAFNSVGDSPVAPPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
Db 181 EYTDWGLLPSEGAFNSVGDSPVAPPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
Qy 241 TPSSGEVLEKAKAQLOQSSNSGKEVLMPSHSLPPASLESSTVTEKSPVLTPTGST 300
Db 241 TPSSGEVLEKAKAQLOQSSNSGKEVLMPSHSLPPASLESSTVTEKSPVLTPTGST 300
Qy 301 EHSIPTPTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLITLPDNEVELKAFVAP 360
Db 301 EHSIPTPTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLITLPDNEVELKAFVAP 360
Qy 361 APPVETTYNENLISHPTDYQGEIKQGHKQTLNLSQLSVGLYVFKVTSSENAFGEFV 420
Db 361 APPVETTYNENLISHPTDYQGEIKQGHKQTLNLSQLSVGLYVFKVTSSENAFGEFV 420
Qy 421 NVTVKPARVNLPPVAVVSPQLOELTLPLTSALDGSQSTDDTEIVSYHWEENGPFTIEE 480
Db 421 NVTVKPARVNLPPVAVVSPQLOELTLPLTSALDGSQSTDDTEIVSYHWEENGPFTIEE 480
Qy 481 KTSVDSPLVRLSNLDPGNSYFRLVTDSGDGATNSTTAALIVNNAVDYPPVANAGNPHIT 540
Db 481 KTSVDSPLVRLSNLDPGNSYFRLVTDSGDGATNSTTAALIVNNAVDYPPVANAGNPHIT 540
Qy 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGEGKHVVMQGVQTPYLHLSAMQEGDYTFOLK 600
Db 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGEGKHVVMQGVQTPYLHLSAMQEGDYTFOLK 600
Qy 601 VTDSSRQOSTAXVTIVOPENNRPPVAVAGDPKELIPFVESATLDGSSSSDDHGIVFYHW 660
Db 601 VTDSSRQOSTAXVTIVOPENNRPPVAVAGDPKELIPFVESATLDGSSSSDDHGIVFYHW 660
Qy 661 EHVGRPSAVENIDKATATVTGLQVGYHFRHTVKKOQGLSSNSTLTAVAKKENNSPPR 720
Db 661 EHVGRPSAVENIDKATATVTGLQVGYHFRHTVKKOQGLSSNSTLTAVAKKENNSPPR 720

Qy 721 ARAGGRHVLVLPNNISITLDGSRSTDDORIYSYLWIRGQSPAAGDVIDGSDHVALQLTN 780
Db 721 ARAGGRHVLVLPNNISITLDGSRSTDDORIYSYLWIRGQSPAAGDVIDGSDHVALQLTN 780
Qy 781 LVEGVYTFHLRVTDQASDSDTATVEVQPPPKSGVLVELTLQVGVGQLTEQRKDTLVQ 840
Db 781 LVEGVYTFHLRVTDQASDSDTATVEVQPPPKSGVLVELTLQVGVGQLTEQRKDTLVQ 840
Qy 841 LAVLLNLDSDIKVQKIRAHSDLSLTIVFYVQSRPPFKVLKAAEVARNLHRLSKEKADF 900
Db 841 LAVLLNLDSDIKVQKIRAHSDLSLTIVFYVQSRPPFKVLKAAEVARNLHRLSKEKADF 900
Qy 901 LIFKVLAVDTAGCLLKCSGHCHCDPLTKRCICSHLWMLNLIORYIWDGESNCESIFVYT 960
Db 901 LIFKVLAVDTAGCLLKCSGHCHCDPLTKRCICSHLWMLNLIORYIWDGESNCESIFVYT 960
Qy 961 VLAFTLVLVTGGFTWLICCCCKRQRTKIRKTKYTITLDNMDQERMELEPKYGIKHRST 1020
Db 961 VLAFTLVLVTGGFTWLICCCCKRQRTKIRKTKYTITLDNMDQERMELEPKYGIKHRST 1020
Qy 1021 EHNSLMVSESEFSDSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSYCKDR 1072
Db 1021 EHNSLMVSESEFSDSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSYCKDR 1072

RESULT 7

ADRO0594

ID ADRO0594 standard; protein; 1072 AA.

XX AC ADR00594;

XX DT 04-NOV-2004 (first entry)

XX DE Human 254p1D6B v.2 protein SEQ ID NO:5.

XX KW 254p1D6B; small interfering RNA; siRNA; immune response;

XX KW 254p1D6B v.2; chromosome 6.

XX OS Homo sapiens.

XX PN WO2004067716-A2.

XX PD 12-AUG-2004.

XX PF 23-JAN-2004; 2004WO-US001965.

XX PR 24-JAN-2003; 2003US-0442526P.

XX XX (AGEN-) AGENSYS INC.

XX PI Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;

XX PI Perez-Villar JJ, Paris M;

XX XX WPI; 2004-580991/56.

XX DR N-ESDB; ADR00593.

XX XX New 254p1D6B siRNA composition comprising a double stranded siRNA that corresponds to the nucleic acid ORF sequence which encodes the 254p1D6B protein or corresponds to a subsequence of the ORF, useful for detecting and treating cancer.

XX PS Claim 1; SEQ ID NO 5; 345pp; English.

XX CC The present invention describes a 254p1D6B small interfering RNA (siRNA) composition that comprises a double stranded siRNA that corresponds to the nucleic acid open reading frame (ORF) sequence which encodes the 254p1D6B protein, or corresponds to a subsequence of the ORF, where the double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides in length. Also described: (1) a composition that comprises, consists essentially of, or consists of a peptide of eight, nine, ten, or eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID

CC NO.3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of Pl; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of Pl; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of Pl; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of Pl; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of Pl. 254P1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254P1D6B v.2, which is used in the exemplification of the present
 CC invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX

SQ Sequence 1072 AA;

Query Match 99.9%; Score 5574; DB 8; Length 1072;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPPTGVLSLLLVTTIAGCARKQCSGRTYNAVISPNETTRIMRVSHPTPVVDCTAA 60
 DB 1 MAPPTGVLSLLLVTTIAGCARKQCSGRTYNAVISPNETTRIMRVSHPTPVVDCTAA 60
 QY 61 CCLSSCDLAWFEGRCYLVSCPHKENCPCPKMGPIRSYLTFLVPVQRPALLDYGDMM 120
 DB 61 CCLSSCDLAWFEGRCYLVSCPHKENCPCPKMGPIRSYLTFLVPVQRPALLDYGDMM 120
 QY 121 LNRGSPGIVGSDPEDIRKDLXFLGKDWGLEEMSEYXDDYRELEKDLQPSGQBPGRSA 180
 DB 121 LNRGSPGIVGSDPEDIRKDLXFLGKDWGLEEMSEYADYRELEKDLQPSGQBPGRSA 180
 QY 181 EYTDWGLLPGSGAFNNSVGSDSPVAPETQDQPELHYLNESASTAPKLPERSVLLPPT 240
 DB 181 EYTDWGLLPGSGAFNNSVGSDSPVAPETQDQPELHYLNESASTAPKLPERSVLLPPT 240
 QY 241 TPSSGEVLEKASQLOEQSSNSGKEVLMPSHSLPPASLELSVTVKESPVLTTPGST 300
 DB 241 TPSSGEVLEKASQLOEQSSNSGKEVLMPSHSLPPASLELSVTVKESPVLTTPGST 300
 QY 301 EHSITPTPTSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
 DB 301 EHSITPTPTSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
 QY 361 APPVETTYNENLISHPTDYQGEIKQGHKOTLNLSQLSVGLYVFKVTVSSNAGFEGFV 420
 DB 361 APPVETTYNENLISHPTDYQGEIKQGHKOTLNLSQLSVGLYVFKVTVSSNAGFEGFV 420
 QY 421 NVTVPARRVNLPPVAVVSPQLQELTLPLTSALIDGSQSTDDTEIVSYHWEERINGPFTIE 480
 DB 421 NVTVPARRVNLPPVAVVSPQLQELTLPLTSALIDGSQSTDDTEIVSYHWEERINGPFTIE 480
 QY 481 KTSVDSVPLRLNLDPGNYSFRLITVTDSDGATNSTTAAALIVNNAVDPYPPVANAGNHTIT 540
 DB 481 KTSVDSVPLRLNLDPGNYSFRLITVTDSDGATNSTTAAALIVNNAVDPYPPVANAGNHTIT 540
 QY 541 LPQNSITLNGQSSDDHQIIVLWESLPGSGEGKHVVMQGVQTPPYLHLSAMQSGDYTFQLK 600
 DB 541 LPQNSITLNGQSSDDHQIIVLWESLPGSGEGKHVVMQGVQTPPYLHLSAMQSGDYTFQLK 600
 QY 601 VTDSSRQQTAXVTTVLVQPPENRPPVAVAGPKDELIFPVESATLDGSSSDDHGVFVHW 660
 DB 601 VTDSSRQQTAXVTTVLVQPPENRPPVAVAGPKDELIFPVESATLDGSSSDDHGVFVHW 660
 QY 661 EHRGFSAVEMENIDKAIATVTGLQVGYTHFRLTVKDDQGLSSTSTLTVAVKENNSPPR 720

DB 661 EHRGFSAVEMENIDKAIATVTGLQVGYTHFRLTVKDDQGLSSTSTLTVAVKENNSPPR 720
 QY 721 ARAGGRHVLPNNISITLDGSRSTDDQRIVSYLWIRDOQSPAAGDIVDGSVHVALQLTN 780
 DB 721 ARAGGRHVLPNNISITLDGSRSTDDQRIVSYLWIRDOQSPAAGDIVDGSVHVALQLTN 780
 QY 781 LVEGVYTHFRLVTDQSGASDSDTATVEVQDPDKSGVLVELTLQVGVGQLTEQRKDTLVRQ 840
 DB 781 LVEGVYTHFRLVTDQSGASDSDTATVEVQDPDKSGVLVELTLQVGVGQLTEQRKDTLVRQ 840
 QY 841 LAVILNVLDSDIKVQKIRAHSDLSITVIVYVQSPPPFKVLAEEVARNLHRLSEKADF 900
 DB 841 LAVILNVLDSDIKVQKIRAHSDLSITVIVYVQSPPPFKVLAEEVARNLHRLSEKADF 900
 QY 901 LLEFKVLVDTAGCLLKSGHGHCDPLTKRCICSHLWMENLIQRYIWDGESNCEWSIFVYT 960
 DB 901 LLEFKVLVDTAGCLLKSGHGHCDPLTKRCICSHLWMENLIQRYIWDGESNCEWSIFVYT 960
 QY 961 VLAFTLIVLTGGFTWLICICCCCKRQKTKIRKTKYTTILDNDMEQERMEIRPKYGIKHRST 1020
 DB 961 VLAFTLIVLTGGFTWLICICCCCKRQKTKIRKTKYTTILDNDMEQERMEIRPKYGIKHRST 1020
 QY 1021 EHNSSLVMSSEFSDSDQDTIFSRKMERGNPKVNMNGSIRNGASFSYCSKOR 1072
 DB 1021 EHNSSLVMSSEFSDSDQDTIFSRKMERGNPKVNMNGSIRNGASFSYCSKOR 1072

RESULT 8
 ADR00592
 ID ADR00592 standard; protein; 1072 AA.
 XX
 AC ADR00592;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Human 254P1D6B v.1 clone LCP-3 protein SEQ ID NO:3.
 XX
 KW 254P1D6B; small interfering RNA; siRNA; immune response;
 KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
 KW 254P1D6B v.1 clone LCP-3; chromosome 6.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 142 /label= unknown
 FT /note= "encoded by MCC"
 FT Misc-difference 157 /label= unknown
 FT /note= "encoded by KCA"
 FT Misc-difference 612 /label= unknown
 FT /note= "encoded by GTR"
 XX
 PN WO2004067716-A2.
 PD
 XD 12-AUG-2004.
 XX
 PF 23-JAN-2004; 2004WO-US001965.
 XX
 PR 24-JAN-2003; 2003US-0442526P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
 PI Perez-Villar JJ, Faris M;
 XX
 DR WPI: 2004-580991/56.
 DR N-PSDB; ADR00591.
 XX
 PT New 254P1D6B siRNA composition comprising a double stranded siRNA that corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B

protein or corresponds to a subsequence of the ORF, useful for detecting and treating cancer.

Claim 1; SEQ ID NO 3; 345pp; English.

The present invention describes a 254p1d6B small interfering RNA (siRNA) composition that comprises a double stranded siRNA that corresponds to the nucleic acid open reading frame (ORF) sequence which encodes the 254p1d6B protein, or corresponds to a subsequence of the ORF, where the double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides in length. Also described: (1) a composition that comprises, consists essentially of, or consists of a peptide of eight, nine, ten, or eleven contiguous amino acids of a protein of figure 2 (Pl, see SEQ ID NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in one of the 42 lists of peptides, given in the specification, or a protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or identical to an entire amino acid sequence of Pl; (2) a polynucleotide that encodes the protein; (3) a composition comprising a polynucleotide that is a full complement of the polynucleotide described above; (4) generating a mammalian immune response directed to the protein of Pl; (5) detecting, in a sample, the presence of a 254p1d6B-related protein or a 254p1d6B-related polynucleotide; (6) a composition that modulates the status of a cell that expresses a protein of Pl; (7) a pharmaceutical composition that comprises the composition described above in a human unit dose form; (8) an antibody or its fragment, which is monoclonal; (9) a non-human transgenic animal that produces the antibody; (10) a hybridoma that produces the antibody; (11) delivering a cytotoxic agent or a diagnostic agent to a cell that expresses the protein of Pl; and (12) inhibiting growth, reproduction or survival of cancer cells that express the protein of Pl. 254p1d6B has cytostatic activity, and can be used in gene therapy. The compositions, molecules and methods are useful for treating and detecting cancer. The present sequence represents the human 254p1d6B v.1 clone LCP-3, which is used in the exemplification of the present invention. The human 254p1d6B gene is located on chromosome 6p22.

Sequence 1072 AA;

Query Match 99.9%; Score 5574; DB 8; Length 1072;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1072; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAPTGVLSLLLVTTAGCARKQCSGRVTSNAVISPNLETRIMRVSHTPFVVDCTAA	60
Db	1	MAPTGVLSLLLVTTAGCARKQCSGRVTSNAVISPNLETRIMRVSHTPFVVDCTAA	60
Qy	61	CCDLSDDLAWFEGRCVLSVCPHKENCEPKKMGPIRSYLTFLVRPQRPQALLDYGDM	120
Db	61	CCDLSDDLAWFEGRCVLSVCPHKENCEPKKMGPIRSYLTFLVRPQRPQALLDYGDM	120
Qy	121	LNRGSPGIWGDSPEDIRKDLXFLGKDWGLEEMSEYDXYRELEKDLIQPSGKQPRGSA	180
Db	121	LNRGSPGIWGDSPEDIRKDLXFLGKDWGLEEMSEYDXYRELEKDLIQPSGKQPRGSA	180
Qy	181	EYTWGLLPGSEGAFNSSVSGSPAVPABTQDDPELHYLNESASTPAPKLPERSVLLPLPT	240
Db	181	EYTWGLLPGSEGAFNSSVSGSPAVPABTQDDPELHYLNESASTPAPKLPERSVLLPLPT	240
Qy	241	TPSSGEVLEKAKAQLOQSSNSGKEVLMPSHSLPPASLELSVTVVEKSPVLTTPGST	300
Db	241	TPSSGEVLEKAKAQLOQSSNSGKEVLMPSHSLPPASLELSVTVVEKSPVLTTPGST	300
Qy	301	EHSIPTPTSAPESTSELPISPTTAPRTVKELTVSAGNLIITLPDNEVELKAFVAP	360
Db	301	EHSIPTPTSAPESTSELPISPTTAPRTVKELTVSAGNLIITLPDNEVELKAFVAP	360
Qy	361	APPVETTYNENWLISHPTDYQGEIKQGHQKQTLNLSQLSVGLYVPKVTVSSENAFGEFV	420
Db	361	APPVETTYNENWLISHPTDYQGEIKQGHQKQTLNLSQLSVGLYVPKVTVSSENAFGEFV	420
Qy	421	NVTVPARRVNLPPVAVVSPQLBELTLPLTSALIDGQSOTDTEIVSYHWEENINGPFIEE	480
Db	421	NVTVPARRVNLPPVAVVSPQLBELTLPLTSALIDGQSOTDTEIVSYHWEENINGPFIEE	480

Qy	481	KTSVDSVPLRLSNLDPGNYSFRLVTDSGATNSTTAALIVNNAVDPYPPVANAGPNHTIT	540
Db	481	KTSVDSVPLRLSNLDPGNYSFRLVTDSGATNSTTAALIVNNAVDPYPPVANAGPNHTIT	540
Qy	541	LPQNSITLNGOSSDDHQIIVLYEWSLGPSEKGVVMQGVQTPYHLHLSAMQEGDYTFOLK	600
Db	541	LPQNSITLNGOSSDDHQIIVLYEWSLGPSEKGVVMQGVQTPYHLHLSAMQEGDYTFOLK	600
Qy	601	VTDSRQOSTAXVTVIIVQPPENRPPVAVAGPKELIPFVESATLDGSSSSDDHGIVFYHW	660
Db	601	VTDSRQOSTAXVTVIIVQPPENRPPVAVAGPKELIPFVESATLDGSSSSDDHGIVFYHW	660
Qy	661	EHVRGSAVEMENIDKAIATVTGLQVGYTHFRLTVKQOGLSSSTLTVAVKKENNSPPR	720
Db	661	EHVRGSAVEMENIDKAIATVTGLQVGYTHFRLTVKQOGLSSSTLTVAVKKENNSPPR	720
Qy	721	ARAGRHVLVLPNNSITLDGSSRSDDORIIVSYLWIRDOQSPAAGDVIDGSDHVALQLTN	780
Db	721	ARAGRHVLVLPNNSITLDGSSRSDDORIIVSYLWIRDOQSPAAGDVIDGSDHVALQLTN	780
Qy	781	LVEGVYTFHLRVTDQSGASDITATVEVQDPDKSLVELTQVGVGQLTQQRKDTLVRQ	840
Db	781	LVEGVYTFHLRVTDQSGASDITATVEVQDPDKSLVELTQVGVGQLTQQRKDTLVRQ	840
Qy	841	LAVLNLVLDSDIKVQKIRAHSDLSITVIVYVQSPRPFPKVLKAAEVARNLHRLSKEKADF	900
Db	841	LAVLNLVLDSDIKVQKIRAHSDLSITVIVYVQSPRPFPKVLKAAEVARNLHRLSKEKADF	900
Qy	901	LLFKVLAVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCEWSIFVYT	960
Db	901	LLFKVLAVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCEWSIFVYT	960
Qy	961	VLAFTLIVLTGGFTWLCTCCCKRQKTKIRKTKYTTILDNMDEQERMELRPKYGIKHRST	1020
Db	961	VLAFTLIVLTGGFTWLCTCCCKRQKTKIRKTKYTTILDNMDEQERMELRPKYGIKHRST	1020
Qy	1021	EHNSLMVSESEFSDQDTIPSRKMERGNPKVSMNGSIRNGASFSYCSKDR	1072
Db	1021	EHNSLMVSESEFSDQDTIPSRKMERGNPKVSMNGSIRNGASFSYCSKDR	1072

RESULT 9
 ADR00862
 ID ADR00862 standard; protein; 1053 AA.
 XX
 AC ADR00862;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Human 254p1d6B v.1 protein sequence SEQ ID NO:273.
 XX
 KW 254p1d6B; small interfering RNA; siRNA; immune response;
 KW 254p1d6B-related protein; cytostatic; gene therapy; cancer; human;
 KW 254p1d6B v.1; chromosome 6.
 XX
 OS Homo sapiens.
 XX
 PN WO2004067716-A2.
 XX
 PD 12-AUG-2004.
 XX
 PF 23-JAN-2004; 2004WO-US001965.
 XX
 PR 24-JAN-2003; 2003US-0442526P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
 PI Perez-Villar JJ, Paris M;
 XX
 DR WPI; 2004-580991/56.
 XX

PT New 254pID6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254pID6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.

PS Example 5; SEQ ID NO 273; 345pp; English.

XX
CC The present invention describes a 254pID6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254pID6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (p1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of p1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of p1; (5)
CC detecting, in a sample, the presence of a 254pID6B-related protein or a
CC 254pID6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of p1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of p1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of p1. 254pID6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents a
CC human 254pID6B v.1 protein sequence, which is used in the exemplification
CC of the present invention. The human 254pID6B gene is located on
CC chromosome 6p22.

XX
SQ Sequence 1053 AA;

Query Match 98.3%; Score 5485; DB 8; Length 1053;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 20 CARQCSGRTYSNAVISNLETRIMRVSHTPVVDCTAACDLSSCDLAWFEGRCYL 79
DB 1 CARQCSGRTYSNAVISNLETRIMRVSHTPVVDCTAACDLSSCDLAWFEGRCYL 60
QY 80 VSCPHKENCEPKWGPPIRSYLTFLVLPVQRPQALLDYGMMLNRGSPSGIWDGSPEDIRK 139
DB 61 VSCPHKENCEPKWGPPIRSYLTFLVLPVQRPQALLDYGMMLNRGSPSGIWDGSPEDIRK 120
QY 140 DLXFLKQWGLEMSYDDYRELEKDLQPSGKQPRGSAEYTDWGLLPGSEGAFNSV 199
DB 121 DLPLFLKQWGLEMSYDDYRELEKDLQPSGKQPRGSAEYTDWGLLPGSEGAFNSV 180
QY 200 GSPAPVPAETQDDPELHYLNESASTPAKLPERSVLLPLPTTSSGCVLEKSKASOLOQ 259
DB 181 GSPAPVPAETQDDPELHYLNESASTPAKLPERSVLLPLPTTSSGCVLEKSKASOLOQ 240
QY 260 SNNSSGKVLMPSHSLPPASLELSVTVVEKSPVLTVTPGSTSHSIPPTPTSAAPSESTPS 319
DB 241 SNNSSGKVLMPSHSLPPASLELSVTVVEKSPVLTVTPGSTSHSIPPTPTSAAPSESTPS 300
QY 320 ELPISPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNYEWNLIISHT 379
DB 301 ELPISPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNYEWNLIISHT 360
QY 380 DYOGELKQGHKQTLNLISQISVLGVYVFKVTVSSNAFEGGVNVTVPARRVNLPPVAVVS 439
DB 361 DYOGELKQGHKQTLNLISQISVLGVYVFKVTVSSNAFEGGVNVTVPARRVNLPPVAVVS 420
QY 440 PQIQELTLPLTSALIDGQSOTDDTEIVSYHWEENGFPIEBKTSVDSVPLRLSLNLDPGNY 499

DB 421 POLQELTLPLTSALIDGQSOTDDTEIVSYHWEENGFPIEBKTSVDSVPLRLSLNLDPGNY 480
QY 500 SFRLTVTDSGATNSTTAAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGNOSDDHQI 559
DB 481 SFRLTVTDSGATNSTTAAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGNOSDDHQI 540
QY 560 VLYEWSLQPGSEGKHVVMQGVQTPVYLHLSAMQEGDYTPQLKVTDSRSRQOSTAXVTIVIQP 619
DB 541 VLYEWSLQPGSEGKHVVMQGVQTPVYLHLSAMQEGDYTPQLKVTDSRSRQOSTAXVTIVIQP 600
QY 620 ENNPPVAVAGPDKELIPFVESATLDGSSSDHGHIVFYHWEHVRGSPSAVEMENIDKAIA 679
DB 601 ENNPPVAVAGPDKELIPFVESATLDGSSSDHGHIVFYHWEHVRGSPSAVEMENIDKAIA 660
QY 680 TVTGLQVCTHYFRLTVKDDQGLSTSTLTAVAKKENNSPPRARAGGRHVLVLPNNSITLD 739
DB 661 TVTGLQVCTHYFRLTVKDDQGLSTSTLTAVAKKENNSPPRARAGGRHVLVLPNNSITLD 720
QY 740 GSRSTDDQRIYSYLWIRDQSPAAAGDVLDGSDHVALQLTNLVEGVYTFHLRVTDSSQAS 799
DB 721 GSRSTDDQRIYSYLWIRDQSPAAAGDVLDGSDHVALQLTNLVEGVYTFHLRVTDSSQAS 780
QY 800 DTDATATVEQPDPRKSGELVELTLQVGVQQLTEQRKDTLVRQLAVLLNVLDSIKVKIKRA 859
DB 781 DTDATATVEQPDPRKSGELVELTLQVGVQQLTEQRKDTLVRQLAVLLNVLDSIKVKIKRA 840
QY 860 HSDLSLTVFVYVQSPRPPEKVLKAAEVARNLHMLSKKADFLFKVLAVDTAGCLLKCSG 919
DB 841 HSDLSLTVFVYVQSPRPPEKVLKAAEVARNLHMLSKKADFLFKVLAVDTAGCLLKCSG 900
QY 920 HGHCDPLTKRCICSHLWMENLIQRYIWDEGNCSEWSIFVYTVLAFTLVLITGFTWLCIC 979
DB 901 HGHCDPLTKRCICSHLWMENLIQRYIWDEGNCSEWSIFVYTVLAFTLVLITGFTWLCIC 960
QY 980 CCKQKQRTKIRKTKYKTYTILDNMDQERMELEPKYGIKHSRSTHNSLWVSESEFSDQDT 1039
DB 961 CCKQKQRTKIRKTKYKTYTILDNMDQERMELEPKYGIKHSRSTHNSLWVSESEFSDQDT 1020
QY 1040 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1072
DB 1021 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1053
RESULT 10
ADR00863
ID ADR00863 standard; protein; 1053 AA.
XX ADR00863;
XX AC ADR00863;
XX 04-NOV-2004 (first entry)
XX Human 254pID6B v.3 protein sequence SEQ ID NO:274.
XX 254pID6B; small interfering RNA; siRNA; immune response;
KW 254pID6B-related protein; cytostatic; gene therapy; cancer; human;
KW 254pID6B v.3; chromosome 6.
XX Homo sapiens.
XX WO2004067716-A2.
XX 12-AUG-2004.
XX 23-JAN-2004; 2004WO-US001965.
XX 24-JAN-2003; 2003US-0442526P.
XX (AGEN-) AGENSYS INC.
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
PI Perez-Villar JJ, Faris M;
XX

DR WPI; 2004-580991/56.
 XX New 254p1D6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254p1D6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.
 XX
 PS Example 5; SEQ ID NO 274; 345pp; English.
 XX
 XX The present invention describes a 254p1D6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254p1D6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254p1D6B-related protein or a
 CC 254p1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254p1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents a
 CC human 254p1D6B v.3 protein sequence, which is used in the exemplification
 CC of the present invention. The human 254p1D6B gene is located on
 CC chromosome 6p22.
 XX
 SQ Sequence 1053 AA;

Query Match 98.3%; Score 5485; DB 8; Length 1053;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 20 CARQCSGRTYSNAVISPNLETRIMRVSHTFPPVDCCTAACCDLSSCDLAWWFEGRCYL 79
 1 CARQCSGRTYSNAVISPNLETRIMRVSHTFPPVDCCTAACCDLSSCDLAWWFEGRCYL 60
 80 VSCPHKENCEPKMGPIRSYLTFLVRPVQRPQALDYGDMMLNRGSPSGIWGDSPEIRK 139
 61 VSCPHKENCEPKMGPIRSYLTFLVRPVQRPQALDYGDMMLNRGSPSGIWGDSPEIRK 120
 140 DLXFLGDKWGLEEMSEYDDYRELEKOLLQPSGQEPGRGSAEYTDWGLLPQSEGAFFNSV 199
 121 DLPLFGDKWGLEEMSEYDDYRELEKOLLQPSGQEPGRGSAEYTDWGLLPQSEGAFFNSV 180
 200 GDSFAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPTTPSSGVELEKASQLEQ 259
 181 GDSFAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPTTPSSGVELEKASQLEQ 240
 260 SSNSGKEVLMPSHSLPPASLELSVTVKESPVLTTPGSTHEHSIPTPTTSAAPSESTPS 319
 241 SSNSGKEVLMPSHSLPPASLELSVTVKESPVLTTPGSTHEHSIPTPTTSAAPSESTPS 300
 320 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNWNLSHPT 379
 301 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNWNLSHPT 360
 380 DYQGEIKQGHKOTLNLSQLSVGLVFKVTVSSNAFGEFGFVNVTKPARRVNLPPVAVS 439
 361 DYQGEIKQGHKOTLNLSQLSVGLVFKVTVSSNAFGEFGFVNVTKPARRVNLPPVAVS 420

Qy 440 POLQELTLPITSALIDGQSQTDDTEIIVSYHWEINGPIBEKTSVDSPLSLNLDPGNY 499
 Db 421 POLQELTLPITSALIDGQSQTDDTEIIVSYHWEINGPIBEKTSVDSPLSLNLDPGNY 480
 Qy 500 SFRLTVTSDGATNSTTAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGNQSSDDHQI 559
 Db 481 SFRLTVTSDGATNSTTAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGNQSSDDHQI 540
 Qy 560 VLYEWSLGPSEGHVVMQVQTPDYHLHLSAQEGDYTFQLKVTDSRQQSQSTAXVTIVQVP 619
 Db 541 VLYEWSLGPSEGHVVMQVQTPDYHLHLSAQEGDYTFQLKVTDSRQQSQSTAXVTIVQVP 600
 Qy 620 ENNRPPVAVAGPKDELIPVESATLDGSSSSDDHGIYFYHWEHVRGSPSAVEMENIDKAI 679
 Db 601 ENNRPPVAVAGPKDELIPVESATLDGSSSSDDHGIYFYHWEHVRGSPSAVEMENIDKAI 660
 Qy 680 TVTGLQVGTTHFRLTVDKQOGLSSTSTLTAVKKNENSPPARAGGRHVLVLPNNSITLD 739
 Db 661 TVTGLQVGTTHFRLTVDKQOGLSSTSTLTAVKKNENSPPARAGGRHVLVLPNNSITLD 720
 Qy 740 GSRSTDQRIIVSVLWIRDGSPAGADVIGSDHSHVALQLTNLVEGVYTFHLRVTDSSQAS 799
 Db 721 GSRSTDQRIIVSVLWIRDGSPAGADVIGSDHSHVALQLTNLVEGVYTFHLRVTDSSQAS 780
 Qy 800 DTDATVEVQDPKPSGLVELTQVGVGQLTEQRKDTLVRLQAVLLNVLDSDIKVKIRA 859
 Db 781 DTDATVEVQDPKPSGLVELTQVGVGQLTEQRKDTLVRLQAVLLNVLDSDIKVKIRA 840
 Qy 860 HSDLSTVIVFVQSRPPPKVLAFAVARNLHMLRSKEKADFLPKVLKRVDTAGCLLKCSG 919
 Db 841 HSDLSTVIVFVQSRPPPKVLAFAVARNLHMLRSKEKADFLPKVLKRVDTAGCLLKCSG 900
 Qy 920 HGHCDPLTKRCICSHLWMENLIQRYIWDGSCNCEWSIFYVTVLAFTLIVLTGGFTWLCIC 979
 Db 901 HGHCDPLTKRCICSHLWMENLIQRYIWDGSCNCEWSIFYVTVLAFTLIVLTGGFTWLCIC 960
 Qy 980 CCKRQKTKIRKTKYITLDNMDEQRMELPKYGIKHRSTEHNSSLMVSESEFSDSDQT 1039
 Db 961 CCKRQKTKIRKTKYITLDNMDEQRMELPKYGIKHRSTEHNSSLMVSESEFSDSDQT 1020
 Qy 1040 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1072
 Db 1021 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1053
 RESULT 11
 ADR00861
 ID ADR00861 standard; protein; 1063 AA.
 XX ADR00861;
 AC ADR00861;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Human 254p1D6B v.3 protein sequence SEQ ID NO:272.
 XX
 KW 254p1D6B; small interfering RNA; siRNA; immune response;
 KW 254p1D6B-related protein; cytostatic; gene therapy; cancer; human;
 KW 254p1D6B v.3; chromosome 6.
 XX
 OS Homo sapiens.
 XX
 FN WO2004067716-A2.
 XX
 PD 12-AUG-2004.
 XX
 PF 23-JAN-2004; 2004WO-US001965.
 XX
 PR 24-JAN-2003; 2003US-0442526P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;

PI Perez-Villar JJ, Faris M;
XX WPI; 2004-580991/56.
XX New 254PID6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX Example 5; SEQ ID NO 272; 345pp; English.
XX The present invention describes a 254PID6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254PID6B-related protein or a
CC 254PID6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254PID6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents a
CC human 254PID6B v.3 protein sequence, which is used in the exemplification
CC of the present invention. The human 254PID6B gene is located on
CC chromosome 6p22.
XX Sequence 1063 AA;
Query Match 98.3%; Score 5485; DB 8; Length 1063;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 20 CARQCSGRTYSNAVISPNLETTIRMVSHTPVVDCTAACDLSLSSCDLAWWFGRCYL 79
DB 11 CARQCSGRTYSNAVISPNLETTIRMVSHTPVVDCTAACDLSLSSCDLAWWFGRCYL 70
QY 80 VSCPHKNCPEKMGPIRSYLFVLRPVORPAQLLDYGDMLNRRGSPSGIWDGSPEDIRK 139
DB 71 VSCPHKNCPEKMGPIRSYLFVLRPVORPAQLLDYGDMLNRRGSPSGIWDGSPEDIRK 130
QY 140 DLXFLGKDWGLEMSYDDYRELEKDLQPSGKQEPGSAEYTDWGLLPGSEGAFNSV 199
DB 131 DLXFLGKDWGLEMSYDDYRELEKDLQPSGKQEPGSAEYTDWGLLPGSEGAFNSV 190
QY 200 GDSVPAVPAETQDPELHYLNESASTPAPKLERSVLLPLPTTPSSGEVLEKEKASQLQDQ 259
DB 191 GDSVPAVPAETQDPELHYLNESASTPAPKLERSVLLPLPTTPSSGEVLEKEKASQLQDQ 250
QY 260 SSSSGKEVLMPSHSLPPASLELSVTVVEKSPVLVTTPGSTSHSIPTPPTSAPSESTPS 319
DB 251 SSSSGKEVLMPSHSLPPASLELSVTVVEKSPVLVTTPGSTSHSIPTPPTSAPSESTPS 310
QY 320 ELPISTTAPRTVKELTVSAGONLIITLPDNEVELKAFVAPAPPVETTYNVEWNLISHTP 379
DB 311 ELPISTTAPRTVKELTVSAGONLIITLPDNEVELKAFVAPAPPVETTYNVEWNLISHTP 370
QY 380 DYQGEIKQGHKQTLNLSQLSVGLYVFKVTSSENAFGEFVNVTVKPARRVNLPPVAVVS 439

DB 371 DYQGEIKQGHKQTLNLSQLSVGLYVFKVTSSENAFGEFVNVTVKPARRVNLPPVAVVS 430
QY 440 POLQELTLPLTSALIDGSSQDDETEIYSYHWEENGPIEBKTSVDSFVLRLSNLDPGNY 499
DB 431 POLQELTLPLTSALIDGSSQDDETEIYSYHWEENGPIEBKTSVDSFVLRLSNLDPGNY 490
QY 500 SFRILTVDSDGATNSTTAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGNQSSDDHQI 559
DB 491 SFRILTVDSDGATNSTTAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGNQSSDDHQI 550
QY 560 VLYEWSLPGSGEGKHVVMQGVQTPYLHLSAMQEGDYTPQLKVTSSRQOSTAXVTVIQOP 619
DB 551 VLYEWSLPGSGEGKHVVMQGVQTPYLHLSAMQEGDYTPQLKVTSSRQOSTAXVTVIQOP 610
QY 620 ENNPPVAVAGPKDELIFPVESATLDGSSDDHGIIFYHWEHVRGSPSAVEMENIDKATA 679
DB 611 ENNPPVAVAGPKDELIFPVESATLDGSSDDHGIIFYHWEHVRGSPSAVEMENIDKATA 670
QY 680 TVTGLQVGYTHFRILTVDKQOGLSSTLTITVAVKKENSSPPRARAGGRHVLVLPNNSITLD 739
DB 671 TVTGLQVGYTHFRILTVDKQOGLSSTLTITVAVKKENSSPPRARAGGRHVLVLPNNSITLD 730
QY 740 GSRSTDDQRIYSYLWIRDPQSPAAAGDVI DGDSDHVALQLTNLVEGVYTFHLRVTDOSQAS 799
DB 731 GSRSTDDQRIYSYLWIRDPQSPAAAGDVI DGDSDHVALQLTNLVEGVYTFHLRVTDOSQAS 790
QY 800 DTDATATVEQDPKRSGLVELTLQVGVQLTEQRKDTLVRQAVLLNVLNLDSDIKVQKIRA 859
DB 791 DTDATATVEQDPKRSGLVELTLQVGVQLTEQRKDTLVRQAVLLNVLNLDSDIKVQKIRA 850
QY 860 HSDLSSTVIVVQSRPPKVLKAAEVARNLHMLLSKEKADFLFKVLKRVDTAGCLLKCSG 919
DB 851 HSDLSSTVIVVQSRPPKVLKAAEVARNLHMLLSKEKADFLFKVLKRVDTAGCLLKCSG 910
QY 920 HGHCDPLTKRCICSHLWMENLIQRYIWDGSCNCSWISFYVTVLAFTLVLITGGFTWLCIC 979
DB 911 HGHCDPLTKRCICSHLWMENLIQRYIWDGSCNCSWISFYVTVLAFTLVLITGGFTWLCIC 970
QY 980 CCKQKRTKIRKTKYITLNDMDQERMLPKYGIKHRSTHNSSLMVSESEFSDQDT 1039
DB 971 CCKQKRTKIRKTKYITLNDMDQERMLPKYGIKHRSTHNSSLMVSESEFSDQDT 1030
QY 1040 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1072
DB 1031 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1063
RESULT 12
ADRO0596
ID ADR00596 standard; protein; 1063 AA.
XX ADR00596;
AC ADR00596;
XX
DT 04-NOV-2004 (first entry)
XX Human 254PID6B v.3 protein SEQ ID NO:7.
DE
XX
XX 254PID6B; small interfering RNA; siRNA; immune response;
KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
KW 254PID6B v.3; chromosome 6.
XX Homo sapiens.
XX WO2004067716-A2.
PN
XX
XX 12-AUG-2004.
XX 23-JAN-2004; 2004WO-US001965.
PF
XX 24-JAN-2003; 2003US-0442526P.
PR
XX (AGEN-) AGENSYS INC.
PA

XX Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
PI Perez-Villar JJ, Paris M;
XX WPI; 2004-580991/56.
DR N-PSDB; ADR00595.
XX New 254p1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254p1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
XX and treating cancer.
XX Claim 1; SEQ ID NO 7; 345pp; English.
XX The present invention describes a 254p1D6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254p1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254p1D6B-related protein or a
CC 254p1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254p1D6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents the
CC human 254p1D6B v.3, which is used in the exemplification of the present
CC invention. The human 254p1D6B gene is located on chromosome 6p22.
XX
SQ Sequence 1063 AA;
Query Match 98.3%; Score 5485; DB 8; Length 1063;
Best Local Similarity 99.7%; Pred No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 20 CARQCSGRTYSNAVISPNLETRIMRVSHTFPVVDCTAACCDLSSCDLAWFEGRCYL 79
DB 11 CARQCSGRTYSNAVISPNLETRIMRVSHTFPVVDCTAACCDLSSCDLAWFEGRCYL 70
QY 80 VSCPHKENCEPKWGPPIRSYLTFLVLRPVRQAQLLDYGDMLNRGSPGIGWGDSPEDIRK 139
DB 71 VSCPHKENCEPKWGPPIRSYLTFLVLRPVRQAQLLDYGDMLNRGSPGIGWGDSPEDIRK 130
QY 140 DLXFLGDKWGLEEMSEYXDDYRELEKDLQPSGQEPKRSABEYTDWGLLPGEAGFNSSV 199
DB 131 DLPLFLGDKWGLEEMSEYSDRELEKDLQPSGQEPKRSABEYTDWGLLPGEAGFNSSV 190
QY 200 GDSAPVPAETQODPELHYLINESASTPAPKLPERSVLLPLPTTPSSGVELEKEKASQLQEQ 259
DB 191 GDSAPVPAETQODPELHYLINESASTPAPKLPERSVLLPLPTTPSSGVELEKEKASQLQEQ 250
QY 260 SSNSGKGVLMPSHSLPPASLELSVTVKEKSPVLVTPGSTEHSIPTPTPTSAAPSESTPS 319
DB 251 SSNSGKGVLMPSHSLPPASLELSVTVKEKSPVLVTPGSTEHSIPTPTPTSAAPSESTPS 310
QY 320 ELPISPTTAPRTVKELTVSAGDNLIITLPDNEVELKAFVAPPVETTYNEMNLISHPT 379
DB 311 ELPISPTTAPRTVKELTVSAGDNLIITLPDNEVELKAFVAPPVETTYNEMNLISHPT 370

QY 380 DYQGEIKQGHKQTLNLSQLSVGLYVFKVTVSSSENAFGGFVNVTVKPARVNLPRVAVVS 439
DB 371 DYQGEIKQGHKQTLNLSQLSVGLYVFKVTVSSSENAFGGFVNVTVKPARVNLPRVAVVS 430
QY 440 PQLQELTLPLTSALIDGSSQSTDDTEIVSYHWEENGPFIEBKTSVDSVLSLNLDRGNY 499
DB 431 PQLQELTLPLTSALIDGSSQSTDDTEIVSYHWEENGPFIEBKTSVDSVLSLNLDRGNY 490
QY 500 SFRLTVTDSGATNSTTAALIVNNAVDPVPPVANAGPNNHTITLPPNSITLNGQSSDDHQI 559
DB 491 SFRLTVTDSGATNSTTAALIVNNAVDPVPPVANAGPNNHTITLPPNSITLNGQSSDDHQI 550
QY 560 VLYEWSLGPSEGGKHVVMQGVQTPYLHLSAMQEGDYTFQLKVTSSROOSTAXVTVIQOP 619
DB 551 VLYEWSLGPSEGGKHVVMQGVQTPYLHLSAMQEGDYTFQLKVTSSROOSTAXVTVIQOP 610
QY 620 ENNRPPVAVAGDPKELIFPVESATLDGSSSSDDHGIIVFYHWEHVRGSPSAVEMENIDKAIA 679
DB 611 ENNRPPVAVAGDPKELIFPVESATLDGSSSSDDHGIIVFYHWEHVRGSPSAVEMENIDKAIA 670
QY 680 TVTGLQVGTYHFRLTVDKQGLSSTLTVAVKKENNSPPARAGGRHVLPNNISITLD 739
DB 671 TVTGLQVGTYHFRLTVDKQGLSSTLTVAVKKENNSPPARAGGRHVLPNNISITLD 730
QY 740 GSRSTDDQRIVSYLWIRDGQSPAAGDVI DGDSDHSAVALQTLNLVVGVTYFHLRVTDGAS 799
DB 731 GSRSTDDQRIVSYLWIRDGQSPAAGDVI DGDSDHSAVALQTLNLVVGVTYFHLRVTDGAS 790
QY 800 DTDATVTEVQDPDKSGELVELTLQVGVQLTQEQKDTLVRQLAVLLNVLDSDIKVQKIRA 859
DB 791 DTDATVTEVQDPDKSGELVELTLQVGVQLTQEQKDTLVRQLAVLLNVLDSDIKVQKIRA 850
QY 860 HSDLSSTVIVFYVQSRPPFKVLKAAEVARNLHMLRSKEKADFLPKVLAVDTAGCLLKCSG 919
DB 851 HSDLSSTVIVFYVQSRPPFKVLKAAEVARNLHMLRSKEKADFLPKVLAVDTAGCLLKCSG 910
QY 920 HGHCDDPLTKRCICSHLWMENLIQRYIWDGSCNCEWSIFVYTVLAFTLIVLTGGFTWLCIC 979
DB 911 HGHCDDPLTKRCICSHLWMENLIQRYIWDGSCNCEWSIFVYTVLAFTLIVLTGGFTWLCIC 970
QY 980 CCKRQKTKIRKTKYTI LNDMDQERWELRPKYGIKHRSTEHNSSLMVSESEFSDQDT 1039
DB 971 CCKRQKTKIRKTKYTI LNDMDQERWELRPKYGIKHRSTEHNSSLMVSESEFSDQDT 1030
QY 1040 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1072
DB 1031 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1063
RESULT 13
ID ADR00599
XX ADR00599 standard; protein; 1063 AA.
AC ADR00599;
XX 04-NOV-2004 (first entry)
DT Human 254p1D6B v.3 protein SEQ ID NO:10.
XX 254p1D6B; small interfering RNA; siRNA; immune response;
KW 254p1D6B-related protein; cytostatic; gene therapy; cancer; human;
XX 254p1D6B v.3; chromosome 6.
OS Homo sapiens.
XX WO2004067716-A2.
XX 12-AUG-2004.
XX 23-JAN-2004; 2004WO-US001965.
XX 24-JAN-2003; 2003US-0442526P.
PR

XX (AGEN-) AGENSYS INC.
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
PI Perez-Villar JJ, Faris M;
XX WPI; 2004-580991/56.
XX New 254p1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254p1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX Example 1; SEQ ID NO 10; 345pp; English.
XX The present invention describes a 254p1D6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254p1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (Pl, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of Pl; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of Pl; (5)
CC detecting, in a sample, the presence of a 254p1D6B-related protein or a
CC 254p1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of Pl; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of Pl; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of Pl. 254p1D6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents the
CC human 254p1D6B v.3, which is used in the exemplification of the present
CC invention. The human 254p1D6B gene is located on chromosome 6p22.
XX Sequence 1063 AA;
SQ Query Match 98.3%; Score 5485; DB 8; Length 1063;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 20 CARQCSGRTYSNAVISNLETTRIMRVSHTFPPVVDCTAACCDLSSCDLAWFPGRCYL 79
Db 11 CARQCSGRTYSNAVISNLETTRIMRVSHTFPPVVDCTAACCDLSSCDLAWFPGRCYL 70
QY 80 VSCPHKENCEPKMGPIRSYLVFLRPVORPAQLLDYGDMLNRGSPSGIWDGSDPEDIRK 139
Db 71 VSCPHKENCEPKMGPIRSYLVFLRPVORPAQLLDYGDMLNRGSPSGIWDGSDPEDIRK 130
QY 140 DLXFLGKDWGLEMSSEYDDYRELEKDLQPSGKQPRGSAEYTDWGLLPGSEGAFNSV 199
Db 131 DLXFLGKDWGLEMSSEYDDYRELEKDLQPSGKQPRGSAEYTDWGLLPGSEGAFNSV 190
QY 200 GDSPPAETQDPELHYLNESASTPAPKLERSVLLPLPTPSSGEVLKESAKOLOBQ 259
Db 191 GDSPPAETQDPELHYLNESASTPAPKLERSVLLPLPTPSSGEVLKESAKOLOBQ 250
QY 260 SNNSSKEVLMPSHSLPPASLELSSVTVKSPVLTVPGSTHSHIPTPTSAAPSESTPS 319
Db 251 SNNSSKEVLMPSHSLPPASLELSSVTVKSPVLTVPGSTHSHIPTPTSAAPSESTPS 310
QY 320 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPPPVETTYNYEWNLISHT 379

Db 311 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPPPVETTYNYEWNLISHT 370
QY 380 DYQGEIKQGHKQTLNLSQLSVGLYVFKVTYSSENAFGEGFVNVTVKPARRVNLPPVAVS 439
Db 371 DYQGEIKQGHKQTLNLSQLSVGLYVFKVTYSSENAFGEGFVNVTVKPARRVNLPPVAVS 430
QY 440 POLQELTPLTSLALIDGSDTDDTEIIVSYHWEENGPFIEEKTSDVSPVLRLSLNDPGNY 499
Db 431 POLQELTPLTSLALIDGSDTDDTEIIVSYHWEENGPFIEEKTSDVSPVLRLSLNDPGNY 490
QY 500 SFRLLTVTSDGATNTTAAALIVNNAVDPYPPVANAGPNHTITLPONSTITLNGNQSSDDHQI 559
Db 491 SFRLLTVTSDGATNTTAAALIVNNAVDPYPPVANAGPNHTITLPONSTITLNGNQSSDDHQI 550
QY 560 VLYEWSLPGSGEGKHVVMQGVQTPYLHLSAMQEGDYTFQLKVTDSRSSQQSTAXVTVIQVP 619
Db 551 VLYEWSLPGSGEGKHVVMQGVQTPYLHLSAMQEGDYTFQLKVTDSRSSQQSTAXVTVIQVP 610
QY 620 ENNRPPVAVAGPDKELIFFVESATLDGSSSSDDHGI VFYHWEHVRGSPSAVEMENIDKAI 679
Db 611 ENNRPPVAVAGPDKELIFFVESATLDGSSSSDDHGI VFYHWEHVRGSPSAVEMENIDKAI 670
QY 680 TVTGLQVGYTHFRLLTVKDOQGLSSTSLTVAVKKENNSPPRARAGRHVLLPNNISITLD 739
Db 671 TVTGLQVGYTHFRLLTVKDOQGLSSTSLTVAVKKENNSPPRARAGRHVLLPNNISITLD 730
QY 740 GSRSTDDQRI VSYLWIRDGQSPAAGDVI DGDSDHSAVALQTLNLVEGVYTFHLRVTDSSQAS 799
Db 731 GSRSTDDQRI VSYLWIRDGQSPAAGDVI DGDSDHSAVALQTLNLVEGVYTFHLRVTDSSQAS 790
QY 800 DTDATATVEVQDPDKRSGELVELTLQVGVQQLTEQQRKOTLVRQLAVLLNVLNVDSDIKVKQIRA 859
Db 791 DTDATATVEVQDPDKRSGELVELTLQVGVQQLTEQQRKOTLVRQLAVLLNVLNVDSDIKVKQIRA 850
QY 860 HSDLSSTVIVFVYVQRRPPFKVLKAAEVARNLHMLRSKEKADPLLKYLAVDVTAGCLLKCSG 919
Db 851 HSDLSSTVIVFVYVQRRPPFKVLKAAEVARNLHMLRSKEKADPLLKYLAVDVTAGCLLKCSG 910
QY 920 HGHCDDLTKRCICSHLWMLNLIQRYIWDGSENCWSIFVYTVLAFTLLVLTTGGFTWLCIC 979
Db 911 HGHCDDLTKRCICSHLWMLNLIQRYIWDGSENCWSIFVYTVLAFTLLVLTTGGFTWLCIC 970
QY 980 CCKQKRTKIRKTKYITILDNMDEQERMELEPKYGIKHRSTHNSLSSVMVSESEFSDQDT 1039
Db 971 CCKQKRTKIRKTKYITILDNMDEQERMELEPKYGIKHRSTHNSLSSVMVSESEFSDQDT 1030
QY 1040 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1072
Db 1031 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1063
RESULT 14
ADGI4994
ID ADGI4994 standard; protein; 978 AA.
XX
AC ADGI4994;
XX
XX 26-FEB-2004 (first entry)
XX
XX Human SECP-27 protein.
XX
KW SECP; secreted protein; neuroprotective; relaxant; antithyroid;
KW antidiabetic; cytosolic; dermatological; immunosuppressive;
KW antiinflammatory; thyromimetic; antiallergic; cerebroprotective;
KW gastrointestinal; hepatotropic; nephrotropic; fungicide; protozoacide;
KW antiparkinsonian; antibacterial; antiparasitic; cardiatic; cardiovascular; anti-HIV;
KW virucide; uropathic; antirheumatic; Parkinson's disease; Alzheimer's; muscular;
KW neurotic; neurodegenerative; catatonia; endocrine; diabetes; leukaemia;
KW myotonic dystrophy; immunological; scleroderma;
KW cervical; breast cancer; immunological; allery; gastrointestinal; renal;
KW systemic lupus erythematosus; allergy; gastrointestinal; parasitic;
KW Goodpasture's syndrome; viral infection; bacterial; fungal; parasitic;
KW protozoal; helminthic; cardiovascular; atherosclerosis; hepatic;

XX 09-JUL-2003; 2003WO-JP008690.
XX 10-JUL-2002; 2002JP-00201856.
XX (TAKE) TAKEDA CHEM IND LTD.
XX Matsuzawa Y, Funahashi T, Shimomura C, Furuyama N;
XX WPI: 2004-122943/12.
XX N-PSDB; ADM36226.
XX Mouse membrane and secretory proteins of adipocyte origin and
XX polynucleotides encoding them for screening compounds as remedies for
XX obesity, diabetes, arteriosclerosis, hypertension and hyperlipemia.
XX Claim 41; SEQ ID NO 22; 195pp; Japanese.
XX This invention relates to novel membrane proteins that are associated
XX with differentiation and/ or metabolic function of adipocytes, in
XX particular of mouse origin. Specifically, it refers to the isolated
XX nucleic acid molecules encoding all or part of these proteins,
XX appropriate antibodies and screening assays useful for the development of
XX drug compositions derived thereof. The present invention describes these
XX compositions as useful for the treatment of diseases associated with
XX abnormalities of adipocyte function, such that they can be used to
XX prevent, treat or diagnose obesity, hypertension, hyperlipaemia, diabetes
XX and arteriosclerosis. Accordingly, they exhibit anorectic, antidiabetic,
XX antiarteriosclerotic, antilipemic and hypotensive activities. This
XX polypeptide is a murine membrane protein sequence related to adipocyte
XX function of the invention.
XX Sequence 1048 AA;
SQ
Query Match 43.5%; Score 2426.5; DB 8; Length 1048;
Best Local Similarity 48.1%; Pred. No. 4.4e-168;
Matches 519; Conservative 136; Mismatches 314; Indels 109; Gaps 16;
QY 10 SLLLVITAGCARKQCEGRYTSNAVISPNLETRIMRYSHTFPVVDCCTAACCDLSSCDL 69
DB 41 SVLWJSTDADESR--CQCKLYGALRTEGE-NHLRLLAGSLUPFHACRAACCRDSACHA 97
QY 70 AMWFEGRYLVSOPHCENCEPKKMGPIRSYLFVLRPVQRPQALLDYDGMMLNRGSPSGI 129
DB 98 LNWLEGWCQFQADCKPQSQCPFRDSSNSML-----IIFQKQTTDD 139
QY 130 WGDSPEDIRKDLXFLKDWGLEB-----MSEYXDYRELEKLLQPSGKQEP 177
DB 140 LGLLPEDDEPHLLRLG--WGRTSWRRQSLGAPLTLSPVSSHQSLRD----- 186
QY 178 GSAEYTDWGLLPGESEAFNSVGDSPAPVAPETQDDPELH-----YLNESASTPAPKLP 230
DB 187 --RQKRDLSVVPFHGAMQHSKVNHSSEAGALSTSAEVRKTTIVAGSFSTNHTTQTPWP 244
QY 231 ERSVLLPLPTTPSSGEGVLEKEKASQIQEQQSSNSGKEVLMPSHSLPPASLELSVTVVEKS 290
DB 245 KNVSIHPEPS-----EHSSPVSG----- 262
QY 291 PVLTVTPGSTHSIPTPPSAAPSES--TPS-ELPISPTAPR-TVKELTVSAGDNLIT 346
DB 263 ---TPQVKSTHSPTDAPLPAVPSYSYATPTPOASSQSTSAPHPVVVKELVWSAGKSVQIT 319
QY 347 LPDNEVELKAFAPAPVETVYNYEWNLISHPTDYQGEIKQGHKQTLNLSQLSVGLYVFK 406
DB 320 LPKNEVQLNAFVLPAPBQETTYTDQQLITHPTDYSGEVERKHSQSLQSLKTPGLYEFK 379
QY 407 VTVSSENAFGEGFVNVTVKPARVNLPPVAVVSPQLQELTLPITSALIDGQSTDDTEIV 466
DB 380 VTVDGQNAHGEVNVTVKPEPKRNPVAVVSPQEQEISLPTTSTIIDGQSTDDDKIV 439
QY 467 SYHWEELNGPFTIEKTSVDSVPLRLNLDPCNYSFRLTVDSGATNSTAALIYNNAVD 526
DB 440 QYHWEELKGPLREEKISDPTAILKLSKLVPGNVTFSLTVDVSDGATNSTASLTVNKAYD 499

QY 527 YPPVANAGPNHTITLPQNSITLNGNQSSDDHQIIVLYEWSLGGSEGHVNVQGVQTPYLH 586
DB 500 YPPVANAGPNQVITLPQNSITLFCNQSTDDHGITSYEWLSLSPSKGKVVNQGVRTPALQ 559
QY 587 LSAMQEGDYTFQLKVTDSRQOSTAKVTIVQVPPNNRPPVAVAGPDKELIIPVESATLDG 646
DB 560 LSAMQEGDYTYQLTVDTTAGQQAQVTVIVQVPPNNKPPQADAGPDKELTLFVDSITLDG 619
QY 647 SSSDDHGIIVFYHWEHVRGSAVEMENIDKAIATVTGLQVGTYHFRLTVKQOQGLSSTST 706
DB 620 SKSTDQRVWSYLWEQSRGPDGVQLENANSSVATVGLQVGTYYFTLVKQERNLQOSS 679
QY 707 LTVAVKKEENSPRARAGGRHVLVLPNNISITLDSRSTDDQRIYSYLWIRDGQSPAAGDV 766
DB 680 VNVIVKEINKPPVAKIAGNVVTVLPTSTAELDGRSSDDKGIYSYLWTRDTSAPAGEV 739
QY 767 IDGSDHVALQLTNLVEGVYTFHLRVTDQSQASDTDTATVEVQDPDKSLVELTLQVGV 826
DB 740 LNHSDHHPVLFLSNLVEGTYTFHLKVTDAKESDTRTVEVKDPDKRSLNVEIILDVNV 799
QY 827 GOLTEQRKDTLAVROLAVLLNVLDSDIIVQKIRAHSDLSLTVIVYVQSRPPPKVLKAAEVA 886
DB 800 SOLTERLKGMLIRQIGVLLGLVDSIIIVQKIQPTQEOSTKMLFFVQNDPPHQLFKGHEVA 859
QY 887 RNLHWRLSKEKADFLFKVLAVDVTAGCLLKCSGHGCDPLTKRCITCSHLMNENLIQRYIW 946
DB 860 AMLKSELQKQKADFLIFALIEISVTVCQLNCSDHGHCDSTFKRCVCDPFWMENFIKVQLR 919
QY 947 DGENCENSWIPYVTVLAPTLLVLTGGFTWLICICCKKQKRTKIRKTKYKTYTILDNMDEQR 1006
DB 920 DGDNSCENSWLYVIIASFVIVVALGILSWTTICCKKQK-GKPKRKRSYKILDATD-QES 977
QY 1007 MELRP--KYGIKHRSTEHNSSLMVSESEFDSDDQDTIFSRKMERGNPKVSMNGSIRNG 1062
DB 978 LELKPTSRAGSKQKGQPTLSSSLMHSELSDD-DAIFTWPDREKGLLYGQNGSVFNG 1034

Search completed: October 12, 2005, 10:02:48
Job time : 131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:00:24 ; Search time 128 Seconds
(without alignments)
4288.663 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 5580
Sequence: 1 MAPPTGVLSLLLVTTIAGC.....VSMNGSIRNGASFSYCSKDR 1072

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result #	Score	Query Match %	Length	DB ID	Description
1	5574	99.9	1109	2	Q9Y4G7
2	4255.5	76.3	1083	2	Q80U39
3	3671	65.8	707	2	Q7L8R6
4	2426.5	43.5	1048	2	Q8K135
5	2403.5	43.1	1085	2	Q8BHU7
6	2403.5	43.1	1085	2	Q8BH23
7	2402	43.0	1049	2	Q8PJ07
8	2401	43.0	1049	2	Q8IZAO
9	2316.5	41.5	946	2	Q8WY25
10	2256.5	40.4	746	2	Q8NDA0
11	2086.5	37.4	691	2	Q9H7V0
12	2004	35.9	639	2	Q86J30
13	1675.5	30.0	603	2	Q8VBZ9
14	1454	26.1	684	2	Q8BHR5
15	1451.5	26.0	1069	2	Q9VSC9
16	1440.5	25.8	685	2	Q9H9L2
17	1390	24.9	985	2	Q7Q8G8
18	1358	24.3	491	2	Q8N2B3
19	1296.5	23.2	640	2	Q8WY39
20	931.5	16.7	353	2	Q96IC3
21	887.5	15.9	695	2	Q7YT23
22	582	10.4	309	2	Q95010
23	452.5	8.1	189	2	Q9BUW6
24	246.5	4.4	1899	2	Q87R00
25	246	4.4	3441	2	Q89PB9
26	238.5	4.3	5020	2	Q8E9W3
27	225	4.0	2009	2	Q7UY44
28	225	4.0	26926	2	Q8WZB3
29	225	4.0	34350	2	Q8WZ42
30	224	4.0	26926	2	Q10466
31	223.5	4.0	2566	2	Q8TSE7

ALIGNMENTS

RESULT 1

Q9Y4G7	PRELIMINARY;	PRT; 1109 AA.
AC	Q9Y4G7; Q9UJC8;	
DT	01-NOV-1999 (TrEMBLrel. 12, Created)	
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	KIAA0319 protein (Fragment).	
GN	Name=KIAA0319;	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
FN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Brain;	
RX	MEDLINE=97349984; PubMed=9205841;	
RA	Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;	
RT	"Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";	
RL	DNA Res. 4:141-150(1997).	
DR	EMBL; AB002317; BAA20777.2; -.	
DR	Genew; HGNC:21580; KIAA0319.	
DR	InterPro; IPR003961; FN III.	
DR	InterPro; IPR000601; PKD.	
DR	SMART; SM00060; FN3; 4.	
DR	SMART; SM00089; PKD; 5.	
DR	PROSITE; PS50093; PKD; 1.	
FT	NON TER 1	
SQ	SEQUENCE 1109 AA; 121961 MW; E39ABFCE6E17BB1C CRC64;	
Query Match 99.9%; Score 5574; DB 2; Length 1109;		
Best Local Similarity 99.7%; Pred. No. 5.8e-265; Mismatches 0; Indels 0; Gaps 0;		
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Qy	1 MAPPTGVLSLLLVTTIAGCARKQCSGRYTSNAVISPNLETTTRIMRVSHFTFPVVDCTAA 60	
Db	38 MAPPTGVLSLLLVTTIAGCARKQCSGRYTSNAVISPNLETTTRIMRVSHFTFPVVDCTAA 97	
Qy	61 CCCLSSCDLAWFEGRCYLVSCHPKNCEPKMGPIRSYLTFLVRPQRPQALLDYGDM 120	
Db	98 CCCLSSCDLAWFEGRCYLVSCHPKNCEPKMGPIRSYLTFLVRPQRPQALLDYGDM 157	
Qy	121 LNRGSPSGIGDSDPEDIRKDLXFLGDKWGLEEMSEYKDDYRELEKDLLQPSGKQEPGSA 180	
Db	158 LNRGSPSGIGDSDPEDIRKDLXFLGDKWGLEEMSEYKDDYRELEKDLLQPSGKQEPGSA 217	
Qy	181 EYTDWGLLPSEGAFFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240	
Db	218 EYTDWGLLPSEGAFFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 277	
Qy	241 TPSSGEVLEKASQLQEQSSNSSGKEVLMPSHSLPPASLESLSSVTVEKSPVLTPTPOST 300	

Db 278 TPSSGEVLEKASQLOEQSSNSGKVLMPSHSLPPASLELSSVTVEKSPVLVTPGST 337
Qy 301 EHSITPTTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLITLDPNEVELKAFVAP 360
Db 338 EHSITPTTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLITLDPNEVELKAFVAP 397
Qy 361 APPVETTYNENWNLISHPDYOGEIKQGHKQTLNLISQLSVGLYVFKVTVSSNAPGEGFV 420
Db 398 APPVETTYNENWNLISHPDYOGEIKQGHKQTLNLISQLSVGLYVFKVTVSSNAPGEGFV 457
Qy 421 NVTVPARRVNLPPVAVVSPQLQELTLPPLTSALIDGSDTDTEIVSYHWEENGPPIEE 480
Db 458 NVTVPARRVNLPPVAVVSPQLQELTLPPLTSALIDGSDTDTEIVSYHWEENGPPIEE 517
Qy 481 KTSVDSVPLRLSLNLDGPNYSFRLTWTSDGATNSTTAALIVNNAVDPVPPVANAGNHIT 540
Db 518 KTSVDSVPLRLSLNLDGPNYSFRLTWTSDGATNSTTAALIVNNAVDPVPPVANAGNHIT 577
Qy 541 LPQNSITLNGQSSDDHQIYLYEWSLPGSGEGKHVVMQGVQTPYLHLGAMQEGDYTFOLK 600
Db 578 LPQNSITLNGQSSDDHQIYLYEWSLPGSGEGKHVVMQGVQTPYLHLGAMQEGDYTFOLK 637
Qy 601 VTDSRQOSTAXVTIVQPNRRPPVAVAGPKELIFPVESATLDGSSSSDDHGIVFYHW 660
Db 638 VTDSRQOSTAXVTIVQPNRRPPVAVAGPKELIFPVESATLDGSSSSDDHGIVFYHW 697
Qy 661 EHVGPASAVEMENIDKATATVGLQVGYHFLTVKDDQGLSSTLTITVAVKKNNSPPR 720
Db 698 EHVGPASAVEMENIDKATATVGLQVGYHFLTVKDDQGLSSTLTITVAVKKNNSPPR 757
Qy 721 ARAGRHVLVLPNNSITLIDGSRSTDQRIVSVLWTRDQSPAAAGDIVDGSNHSVALQLTN 780
Db 758 ARAGRHVLVLPNNSITLIDGSRSTDQRIVSVLWTRDQSPAAAGDIVDGSNHSVALQLTN 817
Qy 781 LVEGYVTFHLRTVDSQASDTPATVEQDPKSGELVELTILQVGVGLTEQRKDTLVRQ 840
Db 818 LVEGYVTFHLRTVDSQASDTPATVEQDPKSGELVELTILQVGVGLTEQRKDTLVRQ 877
Qy 841 LAVILNVLDSDIKVKIRAHSDLSTVIVFYVQSRPPFKVLAARVNLHMLSEKADP 900
Db 878 LAVILNVLDSDIKVKIRAHSDLSTVIVFYVQSRPPFKVLAARVNLHMLSEKADP 937
Qy 901 LFPKVLVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIORYIWDGESNCSEWIFVYT 960
Db 938 LFPKVLVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIORYIWDGESNCSEWIFVYT 997
Qy 961 VLAFTLIVLTGGFTWLCICCCCKRQRTKIRKTKVTYILDNMDEQERMELRPKYGIKHRST 1020
Db 998 VLAFTLIVLTGGFTWLCICCCCKRQRTKIRKTKVTYILDNMDEQERMELRPKYGIKHRST 1057
Qy 1021 EHNSSLMWSESEFSDQDTIFSRKMERGNPKVMNGSIRNGASFSYCSKOR 1072
Db 1058 EHNSSLMWSESEFSDQDTIFSRKMERGNPKVMNGSIRNGASFSYCSKOR 1109

RESULT 2
Q80U39 PRELIMINARY; PRT; 1083 AA.
AC Q80U39;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE MKIAA0319 protein (Fragment).
GN Name=D130043K22Rik; Synonym=mkIAA0319;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;

RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
DR EMBL; AK122446; BAC65528.1; -.
DR MGD; MGI:3036268; DI30043K22Rik.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR000601; PKD.
DR SMART; SM00060; FN3; 4.
DR SMART; SM00089; PKD; 5.
DR PROSITE; PS00093; PKD; 2.
FT NON TER 1
SQ SEQUENCE 1083 AA; 118201 MW; 789BBFEB3A7C02CE CRC64;
Query Match 76.3%; Score 4255.5; DB 2; Length 1083;
Best Local Similarity 75.8%; Pred. No. 2.5e-200;
Matches 820; Conservative 91; Mismatches 160; Indels 11; Gaps 4;
Qy 1 MAPTGVLSLLLVTTAGCARKOCSEGRYTSNAVISNLETTRIMRYSHTPPVVDTAA 60
Db 3 MVSPGVLSSLLLAAMAGSSQCSGRYTSNAVISNLETTRIMRYSHTPPVVDTAA 62
Qy 61 CCDLSSCDLAWFEGRCVLSVCPHKENCEPKMGPIRSYLTFLVRPVPQALLDYGDMM 120
Db 63 CCDLTCDLAWFEGRCVLSVCPHKENCEPKMGPIRSYLTFLVRPVPQALLDYGDMM 122
Qy 121 LNRGSPGWDSPEDIRKOLXFLGKOWGLEMESEYXDYRELEKDLLOPSGQBPGRSA 180
Db 123 LSRGSPGWDSPEDIRKOLXFLGKOWGLEMESEYXDYRELEKDLLOPSGQBPGRSA 182
Qy 181 EYTDWGLLPSCGAFN-SSVCDSPAVPAETQDDP-----ELHYLNESASTPAKLP 231
Db 183 EYDWSLLSPNEGAFNATATGNSAAGMEKIQDTPHPLDQEQALNESTWSPFGHSS 242
Qy 232 RSVLLPLPTTP-SGGEVLEKASQLOEQSSNSGKVLMPSHSLPPASLELSSVTVEKS 290
Db 243 LSSVWPSSASPLPTEEGLEGETLQEQSPNSGKEVPMPSHPSPASLESSPATTEKN 302
Qy 291 PVLVTPTGSEHSITPTTSAAPSESTSELPISPTTAPRTVKELTVSAGDNLITLDPN 350
Db 303 SNFTVTPSRKHSPTTPTSTVLGTLPPLSP-TASRTVKALAVSAGDNLVLTLPDR 361
Qy 351 EVELKAFVAPAPPVETTYNENWNLISHPDYOGEIKQGHKQTLNLISQLSVGLYVFKVTVS 410
Db 362 EAEKASVEPAPPADTTYSYEWLSMHPVDFQGIQENKPTLHLSQLSVGLYAFRAVS 421
Qy 411 SENAFGEFVNVTVKPARVNLPPVAVVSPQLQELTLPPLTSALIDGSDTDTEIVSYHW 470
Db 422 SENAFGEFVNVTVKPARVNLPPVAVVSPQLQELTLPPLTSALIDGSDTDTEIVSYHW 481
Qy 471 BEINGPPIEEKTSVDSPVRLSLNLDGPNYSFRLTWTSDGATNSTTAALIVNNAVDPV 530
Db 482 EVDGPFLEGEFPADTPIRLSLNLDGPNYSFRLTWTSDGATNSTTAALIVNNAVDPV 541
Qy 531 ANAGPNHTITLPONSITLNGQSSDDHQIYLYEWSLPGSGEGKHVVMQGVQTPYLHLGAM 590
Db 542 ANAGPNHTITLPONSITLNGQSSDDHQIYLYEWSLPGSGEGKHVVMQGVQTPYLHLGAM 601
Qy 591 QEGDYTFOLKVTDSRQOSTAXVTIVQPNRRPPVAVAGPKELIFPVESATLDGSSSS 650
Db 602 QEGDYTFOLKVTDSRQOSTAXVTIVQPNRRPPVAVAGPKELIFPVESATLDGSSSS 661
Qy 651 DDHGIVFYHWEHVRGSPASAVEMENIDKATATVGLQVGYHFLTVKDDQGLSSTLTIVA 710
Db 662 DDHGIVFYHWEHVRGSPASAVEMENIDKATATVGLQVGYHFLTVKDDQGLSSTLTIVA 721
Qy 711 VKKNNSPPRACGRHVLVLPNNSITLIDGSRSTDQRIVSVLWTRDQSPAAAGDIVDGS 770
Db 722 VKKNNSPPRACGRHVLVLPNNSITLIDGSRSTDQRIVSVLWTRDQSPAAAGDIVDGS 781

```
Qy 771 DHSVALQTLNVEGVYTHLRVDSOGASDQDTATVEQDPKSGGLVELTLQVGVQLT 830
Db 782 DHRAALQTLNVEGVYTHLRVDSOGASDQDTATVEQDPKSGGLVELTLQVGVQLT 841
Qy 831 EQRKDTLVRLQAVLLNVLDSQKIRAHSDLSSTVIVFYQSRPPPKVLKAAEVAARNLH 890
Db 842 EQKXETLVRLQAVLLNVLDSQKIRAHSDLSSTVIVFYQSRPPPKVLKAAEVAARNLH 901
Qy 891 MRLSKEKADFLFKVLRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESS 950
Db 902 KRLSKEKEAFLFKVLRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESS 961
Qy 951 NCEWSIFVTVLAFLLVLTGFTWLCICCCCKROKRTKIRKTKYTTILDNDQERMELR 1010
Db 962 NCEWSIFVTVLAFLLVLTGFTWLCICCCCKROKRTKIRKTKYTTILDNDQERMELR 1021
Qy 1011 PKYGIKHSRSTHNSLMVSEFSDQDTIFSRERKMERGNPKVSMNGSIRNGASFYSYCSK 1070
Db 1022 PKYGIKHSRSTHNSLMVSEFSDQDTIFSRERKMERGNPKVSMNGSIRNGASFYSYCSK 1081
Qy 1071 DR 1072
Db 1082 DR 1083

RESULT 3
Q7L8R6
ID Q7L8R6 PRELIMINARY; PRT; 707 AA.
AC Q7L8R6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE D073M23.3 (KIAA0319 gene product) (Fragment).
GN Name=D073M23.3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX (1)
RP SEQUENCE FROM N.A.
RA Thorpe K.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031230; CAZ0249.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000601; PKD.
DR SMART; SM00089; PKD; 4.
DR PROSITE; PS50093; PKD; 1.
FT NON TER 1
SQ SEQUENCE 707 AA; 78402 MW; FC7D2A98D778DF46 CRC64;

Query Match 65.8%; Score 3671; DB 2; Length 707;
Best Local Similarity 99.9%; Pred. No. 6.4e-172;
Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 366 TTYNENWNLISHPTDYQGEIKQGHKQTLNLSQLSVGLVYFVKVTVSSNAFEGFVNTVK 425
Db 1 TTYNENWNLISHPTDYQGEIKQGHKQTLNLSQLSVGLVYFVKVTVSSNAFEGFVNTVK 60
Qy 426 PARRVNLPPVAVSPQLQELPLTSLALIDGQSDTDEIIVSYHWEINGPPIEKTSD 485
Db 61 PARRVNLPPVAVSPQLQELPLTSLALIDGQSDTDEIIVSYHWEINGPPIEKTSD 120
Qy 486 SPVLRLNLDGPNYSFRLTVTDSGATNSTTAALIVNADVPPVANAGPNTHTLPQNS 545
Db 121 SPVLRLNLDGPNYSFRLTVTDSGATNSTTAALIVNADVPPVANAGPNTHTLPQNS 180
Qy 546 ITLNGNSSDDHQVTLVLEWSLPGSGKHVVMQGVQTPYLHLSAQMGEDYTFOLKVTDS 605
Db 181 ITLNGNSSDDHQVTLVLEWSLPGSGKHVVMQGVQTPYLHLSAQMGEDYTFOLKVTDS 240
```

```
Qy 606 ROOSTAXVTIVOPENNRPPVAVAGDPKELI FPVESATLDGSSSSDDHGIVFYHWEVRG 665
Db 241 ROOSTAXVTIVOPENNRPPVAVAGDPKELI FPVESATLDGSSSSDDHGIVFYHWEVRG 300
Qy 666 PSVAVENIDKATATVTGLQVGYTHLRVDSOGASDQDTATVEQDPKSGGLVELTLQVGVQLT 725
Db 301 PSVAVENIDKATATVTGLQVGYTHLRVDSOGASDQDTATVEQDPKSGGLVELTLQVGVQLT 360
Qy 726 RHVLVLPNNSITLDGSRSTDDQRIIVSYLWTRDQGS PAAGDIVDGSNHSVALQLTNLVEGV 785
Db 361 RHVLVLPNNSITLDGSRSTDDQRIIVSYLWTRDQGS PAAGDIVDGSNHSVALQLTNLVEGV 420
Qy 786 YTHLRVTVDSOGASDQDTATVEQDPKSGGLVELTLQVGVQLTQTEQRKDTLVRQLAVLL 845
Db 421 YTHLRVTVDSOGASDQDTATVEQDPKSGGLVELTLQVGVQLTQTEQRKDTLVRQLAVLL 480
Qy 846 NVLDSDIKVQKIRAHSDLSSTVIVFYQSRPPPKVLKAAEVAARNLHMLSKKADFLFKV 905
Db 481 NVLDSDIKVQKIRAHSDLSSTVIVFYQSRPPPKVLKAAEVAARNLHMLSKKADFLFKV 540
Qy 906 LRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESSNCEWSIFYTVTLAFT 965
Db 541 LRVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESSNCEWSIFYTVTLAFT 600
Qy 966 LIVLTGFTWLCICCCCKROKRTKIRKTKYTTILDNDQERMELRPKYGIKHSRSTHNS 1025
Db 601 LIVLTGFTWLCICCCCKROKRTKIRKTKYTTILDNDQERMELRPKYGIKHSRSTHNS 660
Qy 1026 LMVSESEFSDQDTIFSRERKMERGNPKVSMNGSIRNGASFYSYCSKDR 1072
Db 661 LMVSESEFSDQDTIFSRERKMERGNPKVSMNGSIRNGASFYSYCSKDR 707

RESULT 4
ID Q8K135 PRELIMINARY; PRT; 1048 AA.
AC Q8K135;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Expressed sequence AU040320.
GN Name=AU040320;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX (1)
RP SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Maman A., Rodriguez S., Sanchez A.,
RA Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Kidney;
```

RA	Strausberg R.;			
RL	Submitted (May-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL;	BC028869;	A4H28869.1;	-
DR	MGI;	2140475;	A0040320.	
DR	InterPro;	IPR003961;	FN III.	
DR	InterPro;	IPR008957;	FN-III-like.	
DR	InterPro;	IPR000601;	PKD.	
DR	SMART;	SM00060;	FN3;	4.
DR	SMART;	SM00089;	PKD;	5.
DR	PROSITE;	PS00993;	PKD;	1.
SEQ	SEQUENCE 1048 AA; 115311 MW; F4D51DD6D84C889D CRC64;			
	Query Match	43.5%;	Score 2426.5;	DB 2; Length 1048;
	Best Local Similarity	48.1%;	Pred. No. 1.1e-110;	
	Matches 519;	Conservative 136;	Mismatches 314;	Indels 109; Gaps 16;
Qy	10	SLLLLVTIAGCARQCSEGRYTSNAVISPNETTRIMESVHTFPVVDCTAACCDLSSCDL	69	
Db	41	SVWLSTDAESR--COQKTLYGAGURTEGE-NHLRLIAGSLPHACRAACCRDSACHA	97	
Qy	70	AWMPEGRCYLVSCPHKENCEPKKMGPIRSYITFVLRPQRPALLDYDGMMLNRPSPGI	129	
Db	98	LWLMGCMGFQADCSKPOSCQPFRTDSSNML-----IIFQKSQTDD	139	
Qy	130	WGSDPEDIRKDLXFLGKDWGLGE-----MSEYXDYRELEKDLLOPSKQKQPR	177	
Db	140	LGLLPEDDEPHLLRLG--WGRTSWRRQSLGAPLTLSPSSHHQSLRD--	186	
Qy	178	GSARYTDWGLLPGSEGAFNSVGSVPAPVPAETQDDPELH-----YLNESASTPAPKLP	230	
Db	187	--ROKRDLSSVVPTHGAMQHSKVNHSSEAGALSPTSAAEVRKKTITVAGSTFSNHTTQTPMP	244	
Qy	231	ERSVLLPLPTTPSPSGEVLKQASQLOQSSNSGKVLMPSHSLPPASLELSSVTVEKS	290	
Db	245	KNVSIHPEPS-----EHSFPGVG-----	262	
Qy	291	PVLTVTPGSTEHSIPTPTPSAAPSES--TPS-ELPISPTTAPR-TVKELTVSAGDNLIIIT	346	
Db	263	--TPQVKSTHSPTDAPLPVAPSVYATPTPQASSQSTSAHPVPVKELVVSAGKSVQIT	319	
Qy	347	LPDNEVELKAPVAPPVETTYNTYEWNLISHTPDYQGIKQHKQTMLNSQLSVGLYVFK	406	
Db	320	LPRNEVLQNAFLVPEAFEGEYTYDWMQILTHPTDYSGEVRKHSQSLQKLTPLGLYEFK	379	
Qy	407	VTVSSENAFGFGVNVTVKPARVNLPPVAVVSPQLQELTPLTSAALIDGQSOTDDTRIV	466	
Db	380	VTVDQNAHGEGYVNVTVKPEKRRNPVAVVSPQFQEISLPTTITIIDGQSOTDDDKIV	439	
Qy	467	SYHWEEINGPPIEBKTSVDSPLRLSNLDPGNYSFRLTVTDSGATNSTTAAIIVNNAVD	526	
Db	440	QYHWEEKGLPREEKISDITALLKLSLVPNGYTFSLTVDSGATNSTTASLTVKNKAVD	499	
Qy	527	YPPVANAGPNHTITLTPQNSITLNGQSSDDHQIVLYEWSLGPSSGKGVVMQGVOTPYLH	586	
Db	500	YPPVANAGPNQVITLTPQNSITLFGNQSTDHIGITSYEMSLSPSSKSGKVVEMQGVATPALQ	559	
Qy	587	LSAMQEGDYYQLKVTDSRSSQSTAKXTVIVQPENNRPPVAVAGDPKELIIPVESATLDG	646	
Db	560	LSAMQEGDYYQLTVTDTAGQQAQVTVIVQPENNKPPQADAGDPKEULTLPVDSITLDG	619	
Qy	647	SSSDDDHGIIVFYHWEHVRGSPAVEMENIDKAIATVTLQVGTYHFRLTVKDOQGLSSTST	706	
Db	620	SKSTDDQRVVSYLWEQSRGPDGVQLLENANSSVATVTGLQVGYTVTLTVKDERNLQSSS	679	
Qy	707	LTVAVKKENNSPPRARAGRHVLPVLPNNSITLDGSRSTDQRIIVSYLYMRDQSSPAAGDV	766	
Db	680	VNVIKKEINKPPVAKIAGNVVVTLPSTAELDGRSSDDKGIIVSYLYMTRDTSPPAAGEV	739	
Qy	767	IDGSDHSVALQITNLVEGYVTFHLRVTSQGNASDPTATVVEQDPKSGVLVELTLQGVV	826	
Db	740	LNHSDHHPVLFLSNLVEGTYTFHLKVTDAKGESDTRTTVVEVKPPRPSNLVETIILDVNV	799	
Qy	827	GQLTEQRKDLTVROLAVLLINVLNLDSDIKVQIKRAHSDLSITVIVFYVQSRPPFPKVLKAAEVA	886	

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Babak S.A., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McWeney P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.D., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M.C., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Smutuz J., Myers R.W., Butterfield Y.S.,
RA Krzywinski A.I., Skalska J., Smaltz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2].
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RL Strauberg R.;
RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014530; AAH14530.1; -.
DR SMART; IPRO00601; PKD.
DR InterPro; PS00089; PKD; 5.
DR PROSITE; PS00093; PKD; 1.
SQ SEQUENCE 1049 AA; 115584 MW; 5BDE6057C09A649 CRC64;

Query Match 43.0%; Score 2402; DB 2; Length 1049;
Best Local Similarity 48.0%; Pred. No. 1.8e-109;
Matches 519; Conservative 134; Mismatches 313; Indels 116; Gaps 17;

Qy 10 SLLLVTTAGCARQCSEGRYSNAVISPNLETRIMRVSHVTPVVDCTAACCDLSSCDL 69
Db 41 SVLWLTDA--SESRCCQKTFQVGLRSGE-NHLMLEGTPLSQSLAACQDSACHV 97

Qy 70 AWPFEGRCYLVSCPHKENCEPKMGPIRSYLTFLVRPVPQPAQLLDYGDMMNRGSPGI 129
Db 98 FWMLEGMCIQADCSRPOSCRAFRTSHSSNSMLVP-LKKFQTADDL----- 140

Qy 130 WGDSPDIRKDLXFLGKDWGLEEMSEYKDDYRELEKDLQPSGKQEPGRGSAEYTDWGLLP 189
Db 141 -GFLPEDDDVPHLLGLGNWA-----SWRQSPRAAL----- 170

Qy 190 GSEGAFFSVGDSAPVAPARTQDDPELHYLNESASTPAPKLPER-----SVLLPLPTPSSG 245
Db 171 -----RAVSSSDQ-----SLIRKQKRGSPSDVVTPIVTOHS-- 204

Qy 246 EVLEKEKASQLEQSSNSGKVLMPGSHLPPASLELS-----SVTVKSPVLTVTPGS 299
Db 205 KYNDSNELGLTTSGSAEVHKAITISSPLTDLTAELSGGPKNVSVQPEISEGLATTFST 264

Qy 300 -----TEHSITPP-----PTSAAPSESTPSELPTSPPTAPRTVKELTVSAGDN 342
Db 265 QQVKSSSEKTAIVAPQVPAPSYATPTPQASFQSTSAFPVP-----IKELVVSAGES 316

Qy 343 LIITLPDNEVELKAFAPAPVETTYNYEWNLISHPTDYQGEIKQGHKQTLNLSQSLVL 402
Db 317 VOITLPKNEVQLNAYVLQEPPEKGTITYTDWQLITHPRDYSGEMEGKSHQILKLSKLTPL 376

Qy 403 YVFKVTVSSNAFGEFNVTVKARRVNLPPVAVSPOLQELTLPITSALIDGQSQTDD 462
Db 377 YEFKVIQVQNAHGEYVNVTVKPEPRKNRPPIALVSPQFQEIPLTTSITVIDGQSQTDD 436

Qy 463 TEIVSYHWEENGPIIEKTSVDGSPVLRLSLNDLPNGYSFRILTVDSDGATNTSTAALVN 522
Db 437 DKIVQYHWEELKPLREKISEDTAILKLSKLVPGNYTFTSLTVVDSGATNTSTAALVN 496

Qy 523 NAVDPPPVANAGPNHTITLPONSITLNGNQSSDDHQIVLYEWSLGPSEGHVVMQGVOT 582
Db 497 KAVDPPPVANAGPNQVITLPQNSITLFGNQSTDDHGITSYEWLSLSPSKGKVMQGVRT 556

Qy 583 PYLHLSAMQEGDYTFOLKVTQSSRQOSTAXVTYVQPNENRPPVAVAGPKDELFPVESA 642
Db 557 PTLQLSAMQEGDYTYQLTVDITIGQATAQVTVIVQPNENRPPQADAGPKDELTPVDST 616

Qy 643 TLDGSSSDDHGIVFYHMEHVRGSPSAVEMENIDKAIATVTGLQVGYHFLTVKDOOGLS 702
Db 617 TLDGSSSDDHGIVFYHMEHVRGSPSAVEMENIDKAIATVTGLQVGYHFLTVKDOOGLS 702

Qy 703 STSLTAVAVKKNNSPPRARAGRHVLVLPNNSTLTDGSRSTDDQRIVSVYLIWIRDGOSPA 762
Db 677 SQSSVNVVKEEINKPPIAKITGNVITLPTSTAEALDGSKSSDDKGVSVLWTRDEGSPA 736

Qy 763 AGVIDGSDHSVALQTLNVEGVYTFHLRVTDGSGASDSTATATVEVQDDPKSGVLVELTL 822
Db 737 AGEVLNHSDDHPLFLFLSNLVEGYTFHLKVTDAGESDTRTVEVQDDPKSGVLVELTL 796

Qy 823 QVGVQLTQTEQRKDTLVRQLAVLLNLDSDIKVKIRAHSDLSSTVIVFVQSRPPFKVLKA 882
Db 797 DINVSQUTERLKGMPFIQIUGVLLGLVSDIIVQIPIYEQSTKMWVFPVQNEPHQIFKG 856

Qy 883 AEVARNLHMLRSKEKADFLFKVLVRVDTAGCLLKCSGHGCDPLTKRCICSHLWMLNIQ 942
Db 857 HEVAAMLKSELKQKADFLIFRALEVNVTVCQLNCSDHGCHDSFTKRCICDPFWMENFIK 916

Qy 943 RYIWDGSDNCESWIFVYTVLAFTLIVLTGGFTWLCICCCRQKQKTKIRKTKYTILDMND 1002
Db 917 VQLRDGSDNCESWVLYYIATFVIWVALGILSWTVICCCRQK-GKPKRSKYKILDATD 975

Qy 1003 EOERMEIRP--KYGIKHSSTHNSSLMVSESEPSDDDTIFSRKEMRGNPKVSMNGSIR 1060
Db 976 -QESLEKPTSRAGIKOKGLLSLHMSSELDSD-DAIPTWPDREKGLLHGQNGSVP 1033

Qy 1061 NG 1062
Db 1034 NG 1035

RESULT 8
Qy Q81ZAO PRELIMINARY; PRT; 1049 AA.
Db Q81ZAO
AC Q81ZAO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polycystic kidney disease 1-related protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY163234; AAN61054.1; -.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR000601; PKD.
DR SMART; SM00060; FN3; 4.
DR SMART; SM00089; PKD; 5.
DR PROSITE; PS00093; PKD; 1.
SQ SEQUENCE 1049 AA; 115717 MW; 1781BFCA3C765DEB CRC64;

Query Match 43.0%; Score 2401; DB 2; Length 1049;
Best Local Similarity 48.0%; Pred. No. 2e-109;
Matches 519; Conservative 134; Mismatches 313; Indels 116; Gaps 17;

Qy 10 SLLLVTTAGCARQCSEGRYSNAVISPNLETRIMRVSHVTPVVDCTAACCDLSSCDL 69
Db 41 SVLWLTDA--SESRCCQKTFQVGLRSGE-NHLMLEGTPLSQSLAACQDSACHV 97

Qy 70 AWPFEGRCYLVSCPHKENCEPKMGPIRSYLTFLVRPVPQPAQLLDYGDMMNRGSPGI 129
Db 98 FWMLEGMCIQADCSRPOSCRAFRTSHSSNSMLVP-LKKFQTADDL----- 140

Qy 130 WGDSPDIRKDLXFLGKDWGLEEMSEYKDDYRELEKDLQPSGKQEPGRGSAEYTDWGLLP 189
Db 141 -GFLPEDDDVPHLLGLGNWA-----SWRQSPRAAL----- 170

Qy 190 GSEGAFFSVGDSAPVAPARTQDDPELHYLNESASTPAPKLPER-----SVLLPLPTPSSG 245
Db 171 -----RAVSSSDQ-----SLIRKQKRGSPSDVVTPIVTOHS-- 204

Qy 246 EVLEKEKASQLEQSSNSGKVLMPGSHLPPASLELS-----SVTVKSPVLTVTPGS 299
Db 205 KYNDSNELGLTTSGSAEVHKAITISSPLTDLTAELSGGPKNVSVQPEISEGLATTFST 264

Qy 300 -----TEHSITPP-----PTSAAPSESTPSELPTSPPTAPRTVKELTVSAGDN 342
Db 265 QQVKSSSEKTAIVAPQVPAPSYATPTPQASFQSTSAFPVP-----IKELVVSAGES 316

Qy 343 LIITLPDNEVELKAFAPAPVETTYNYEWNLISHPTDYQGEIKQGHKQTLNLSQSLVL 402
Db 317 VOITLPKNEVQLNAYVLQEPPEKGTITYTDWQLITHPRDYSGEMEGKSHQILKLSKLTPL 376

Qy 403 YVFKVTVSSNAFGEFNVTVKARRVNLPPVAVSPOLQELTLPITSALIDGQSQTDD 462
Db 377 YEFKVIQVQNAHGEYVNVTVKPEPRKNRPPIALVSPQFQEIPLTTSITVIDGQSQTDD 436

Qy 463 TEIVSYHWEENGPIIEKTSVDGSPVLRLSLNDLPNGYSFRILTVDSDGATNTSTAALVN 522
Db 437 DKIVQYHWEELKPLREKISEDTAILKLSKLVPGNYTFTSLTVVDSGATNTSTAALVN 496

Qy 523 NAVDPPPVANAGPNHTITLPONSITLNGNQSSDDHQIVLYEWSLGPSEGHVVMQGVOT 582
Db 497 KAVDPPPVANAGPNQVITLPQNSITLFGNQSTDDHGITSYEWLSLSPSKGKVMQGVRT 556

Qy 583 PYLHLSAMQEGDYTFOLKVTQSSRQOSTAXVTYVQPNENRPPVAVAGPKDELFPVESA 642
Db 557 PTLQLSAMQEGDYTYQLTVDITIGQATAQVTVIVQPNENRPPQADAGPKDELTPVDST 616


```
Db 755 EVAAMKSELKQKADFLIFRALEVNTVTCQLNCSHGHCDSFTKRCICDPLFWMFNFKV 814
Qy 944 YWPGESNCSEWSIFVTVVLAFTLVLTGFTWLCICCCCKRQRTKIRKTKYITILDNMDE 1003
Db 815 QLRDGSNCSEWSLVYIATFVIVVALGILSWTVICCCCKRQK-GKPKRKSXYKILDATD- 872
Qy 1004 QERMELRP--KYGIKHSRTEHNSLSMVSESEFSDSDODTIFREKMERGNPKVSMNGSIRN 1061
Db 873 QESLEKPTSRAGIKQKGLLSLSMVSESELDSD-DAIFTWPDREKGLLHGQNGSVFN 931
Qy 1062 G 1062
Db 932 G 932

RESULT 10
QNDAAO
ID Q8NDAO PRELIMINARY; PRT; 746 AA.
AC Q8NDAO;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp434C0829 (Fragment).
GN Name=DKFZp434C0829;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834315; CAD38985.1; -
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR00601; PKD.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00089; PKD; 5.
DR PROSITE; PS50093; PKD; 1.
KW Hypothetical protein.
SQ
NON TER 1
SEQUENCE 746 AA; 82461 MW; A0945E57D8243191 CRC64;

Query Match 40.4%; Score 2256.5; DB 2; Length 746;
Best Local Similarity 60.2%; Pred. No. 1.6e-102;
Matches 441; Conservative 98; Mismatches 189; Indels 5; Gaps 4;

Qy 332 VKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNYEWNLSHPDYOGEIKQGHKQ 391
Db 3 IKELVVSAGESVQITLPKNEVQLNAYLQEPKGETTYTDQLITHPDYSGEKGSQ 62
Qy 392 TLNLSQSLVGLYFVKTVSSNAFGEFVNVTVPARRVNLPPVAVSPQLQLPLPTS 451
Db 63 ILKLSKLPGLYEFKVIVEGQNAHGEVYVTVKPEPKRNPPIAIVSPQFQELSLPTS 122
Qy 452 ALIDGSQSTDTEIVSYWEEINGPFIEKTSVDSPLRLNLDPCNYSPLTIVTDSGA 511
Db 123 TVIDGSQSTDDBKIVQYHWEELKGPLREEKISDAILKLSKLVPGNYSPLTIVTDSGA 182
Qy 512 TNSTTAALIVNNAVDPYPPVANAGPNHHTITLPNSITLNGNOSDDHQVLVWLSLPGOSE 571
Db 183 TNSTTANITVNAVDPYPPVANAGPNQVITLPNSITLFGNSTDDHGTSTWLSLSPSK 242
Qy 572 GKHVVMQGVQTPYPLHLSAMQEGDYTFOLKVTDSRQSQSTAXVTIVQPPENRPPVAVAGP 631
Db 243 GKVEMQGVRTPTLQLSAMQEGDYTYQLTVDITIGQATAQVTVIVQPPENKPPQADAGP 302
Qy 632 DKELIFPVESATLDGSSSDHGIVFYVFWHVRGPSAVEMENIDKAITVTLQVGTTHF 691
```

```
Db 303 DKELTLPVDSTTLDSGSKSDQKIISLWEKTQCPGCVQLENANSSVATVTLQVGTVP 362
Qy 692 RLTVKDDQGLSSSTLTVAVKKENNSPPRACGRHVLVLPNNSITLDGSRSTDDQRIYS 751
Db 363 TLTVDKERNLQSSSVNVIVKEINKPPIAKITGNVITLPTSTAELDSGSKSDKGIVS 422
Qy 752 YLWTRDGSQSPAAGDIVDSDHSHVALQLTNLVEGVYTFHLRVTDQSGASDPTATVEQPD 811
Db 423 YLWTRDGSQSPAAGEVNLHSDHHPILFNLVEGVYTFHLKVTDAKGSDDTDTTVEVKPD 482
Qy 812 PRKSLVELTLQVGVGQLTQEQKDTLVRLQALVNLVLDSDIKVQKIRAHSDLSVIFYV 871
Db 483 PRKNLVEIILDINVSQLTRELKGMFIRQIGVLLGVLDSDIIVQKIQPYTEQSTKMWFFV 542
Qy 872 QSRPPFKVLKAAEVARNLHMLRSLSEKADFLFKVLRYDVTAGCLLKCSGHGCHDPLTKRCI 931
Db 543 QNEPFPQIFKGEHVAAMLKSELKQKADFLIFRALEVNTVTCQLNCSHGHCDSFTKRCI 602
Qy 932 CSHLWMENLIQRYIWDGESNCSEWSIFVTVVLAFTLVLTGFTWLCICCCCKRQRTKIRK 991
Db 603 CDPFWMENFIKVQRDGSNCSEWSLVYIATFVIVVALGILSWTVICCCCKRQK-GKPKR 661
Qy 992 KTKYTILDNDDEQERMELRP--KYGIKHSRTEHNSLSMVSESEFSDSDODTIFREKMERG 1049
Db 662 KSKYKILDATD-QESLEKPTSRAGIKQKGLLSLSMVSESELDSD-DAIFTWPDREKG 719
Qy 1050 NPKVSMNGSIRNG 1062
Db 720 KLLHGQNGSVPNG 732

RESULT 11
Q9H7V0
ID Q9H7V0 PRELIMINARY; PRT; 691 AA.
AC Q9H7V0;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein FLJ14225 (PKD1-like protein).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Saito K., Kawai Y., Isono Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirooka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Horita T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Mutsaers K., Yuuki H., Oshima A., Sasaki N., Aotaka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo K., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
```

Query Match	37.4%	Score 2086.5	DB 2	Length 691
Best Local Similarity	60.7%	Pred. No. 3.1e-94		
Matches 410	Conservative 88	Mismatches 173	Indels 5	Gaps 4
Qy 389	HKQTLNLSQLSVGLYKVFVTVSSNAFGEFVNVTVKPARRVNLPPVAVVSPQLQELTLP	448		
Db 5	HSQILKLSKLTPLGYEFKVIQVGHGEGVYVTVKPEPRKRPPIAIVSPQFQEISLP	64		
Qy 449	LTSALIDGSGTDDTEIIVSYHWEELNGPFIEKTSVDFVLNSLNDGNTVFRLLTVTDS	508		
Db 65	TTSTVIDGSGTDDDKIVQYTHWEELKGLPREEKISEDTAILKLSKLVPGNVTFTSVTVDS	124		
Qy 509	DCATNSTTAALIVNNAVDPVPPVNAAGPNHTITLPNSTITLNGNSQSDHDIIVLYEWSLCP	568		
Db 125	DCATNSTTANITVNAVDYPPVNAAGPNQVITLPQNSITLFGNQSTDDHGITSYEWLSLP	184		
Qy 569	GSEGVVVMQGVQTPPYLHLSAQEGDYTFOLKVTDSRRQSTAXTVTVVQPPNNRPVAV	628		
Db 185	SSKGVVMQGVQVPTPLQLSAMQEGDYTVQLTVTDTIGQATAQTVTVVQPPNNKPPQAD	244		
Qy 629	AGPKELIPVESATLDGSSSSDDHGIVFYVHWEHVGSGVAVEMENIDKAIATVTGLQVGT	688		
Db 245	AGPKELTLPVDTSTLTDGSSSSDDKIIISYLWEKTCQGPDGVLNANSGVATVTGLQVGT	304		
Qy 689	YHRLTVKDCQGLSTSTLTVAVKKENNSPPRARAGGRHLVLPNNSITLDCGSRSTDDQR	748		
Db 305	YVFTLTVDERNLQSSQSVNIVKEINKPPIAKITGNVVTITLPITSTALDGSKSSDDKG	364		
Qy 749	IVSYLWIRDGSGPAAGDVIDGSDHSVALQLTNLVEGVYTFHLRVTVDSQASDITATVVEV	808		

Db 186 SSKGVVEMQVTRPTLQLSAMQSGDYTYQLTVTDITGQQATAQVTVIVPENNKPQAD 245
Qy 629 AGDPKELIFPVESATLDGSSSDHGIYFVHWEHVRGSAVEMENIDKAIATVTVGLQVGT 688
Db 246 AGDPKELTLPVDSITLDGSSSDDKIISYLWERTQDGDVQLENANSSVATVTVGLQVGT 305
Qy 689 YHFLRTVQDQGLSTSTLTAVKKNNSPPRAGRGHVLVLPNNSTLIDGSRSTDDOR 748
Db 306 YVFLTVKDERNLQSSQSVNVIVKEINKPPIAKITGNVITLTSTAEALDGSXSDKG 365
Qy 749 IVSYLWIRIDGQSPAAGVDIDGSDHVALQLTNLVEGYVTFHLRVTDGSGASDITDTATVEV 808
Db 366 IVSYLWIRIDGSPAAGVNLHSDHHPFLFLSNLVEGYTFHLRVTDGSGASDITDTATVEV 425
Qy 809 QPDRKSGELVELTQVQGLTEQKOTLVRQLAVLNLVDSIDKVKQIRAHSDLSTVIV 868
Db 426 KPDRKKNLVEIILIDINVSQLTERLKGFIQIGVLLGLVDSIDIIVQKIQPTQSTQWV 485
Qy 869 FYQSRPPFKVLKAAEVARNLHMLSKKADFLPKVLRVDTAGCLLKCSGHGCHDPLTK 928
Db 486 FFVQNEPPOHIFKGHEVAAMLKSELKQKADFLIFRALEVNVTVCQLNCSDHGCHDSFTK 545
Qy 929 RCICSHLWENLIQRYIWDGSENCSEWSIFVYTVLAFTLIVLTGFTWLCICCCCKRQRTK 988
Db 546 RCICDPFWMENFIKVLQDRGDSNCSEWSVLYIIATFVIVVALGILSVTIVCCCKRQK-GK 604
Qy 989 IRKTKYKTLIDNMQERMLRP 1011
Db 605 PKRSKYKILDATD-QESLELKP 626

RESULT 13
ID Q8VBZ9 PRELIMINARY; PRT; 603 AA.
AC Q8VBZ9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AU040320 protein (Fragment).
GN Name=AU040320;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC022154; AAH22154.1; -.
DR MGD; MGI:2140475; AU040320.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00089; PKD; 3.
DR PROSITE; PS00093; PKD; 1.
FT NON_TER 1
SQ SEQUENCE 603 AA; 65967 MW; BE0DA1665D4046BD CRC64;

Query Match 30.0%; Score 1675.5; DB 2; Length 603;
Best Local Similarity 59.4%; Pred. No. 3.6e-74;
Matches 328; Conservative 76; Mismatches 143; Indels 5; Gaps 4;

Qy 513 NSTTAALIVNADVPPVAVAGNPHHTITLPNSITLNGNSSDDHQIVLYEWSLGPSEG 572
Db 41 NSTTASLTAVKADVPPVAVAGNPHHTITLPNSITLFGNQSTDDHGITSYEWSLSPSKG 100
Qy 573 KHVVMQGVQTPYLHLSAMQSGDYTFQLKVTDSRQOSTAXVTVIVORENNRPPVAVAGPD 632
Db 101 KVVEMQGVTRPALQLSAMQSGDYTYQLTVTDITGQQATAQVTVIVPENNKPQADAGPD 160
Qy 633 KELIFPVESATLDGSSSDHGIYFVHWEHVRGSAVEMENIDKAIATVTVGLQVTVHFR 692
Db 161 KELTLPVDSITLDGSKSTDQORVVSYLWESRGDPGVQLENANSSVATVTVGLQVTVYFT 220
Qy 693 LTVKDDQGLSTSTLTAVKKNNSPPRAGRGHVLVLPNNSTLIDGSRSTDDQRIVSY 752
Db 221 LTVKDERNLQSSQSVNVIVKEINKPPIAKITGNVITLTSTAEALDGSRSDDKGIYVY 280
Qy 753 LWTIDGQSPAAGVDIDGSDHVALQLTNLVEGYVTFHLRVTDGSGASDITDTATVEVQDP 812
Db 281 LWTIDGQSPAAGVNLHSDHHPFLFLSNLVEGYTFHLRVTDGSGASDITDTATVEVQDP 340
Qy 813 RKSGLVELTQVQGLTEQKOTLVRQLAVLNLVDSIDKVKQIRAHSDLSTVIVFVQ 872
Db 341 RKSNLVEIILIDNVSQLTERLKGFIQIGVLLGLVDSIDIIVQKIQPTQSTQWV 400
Qy 873 SRPEPKVLKAAEVARNLHMLSKKADFLPKVLRVDTAGCLLKCSGHGCHDPLTKRQK 932
Db 401 NDPPHQLFKGHEVAAMLKSELKQKADFLIFRALEVNVTVCQLNCSDHGCHDSFTKRCVC 460
Qy 933 SHLWENLIQRYIWDGSENCSEWSIFVYTVLAFTLIVLTGFTWLCICCCCKRQRTK 992
Db 461 DPFWMENFIKVLQDRGDSNCSEWSVLYIIATFVIVVALGILSVTIVCCCKRQK-GKPKRK 519
Qy 993 TKYTLIDNMQERMLRP--KYGIKHSSTHNSLAVSRSEFSDSDDTIFSRKMERGN 1050
Db 520 SRYKILDATD-QESLELKPTRAGSKOKGPTLSSLMHSELSLSD-DALFTWDPREKQK 577
Qy 1051 PKVSMNGSIRNG 1062
Db 578 LLYQGQSVENG 589

RESULT 14
Q8BHR5 PRELIMINARY; PRT; 684 AA.
ID Q8BHR5
AC Q8BHR5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DB Mus musculus 7 days embryo whole body cDNA, RIKEN full-length enriched
DB library, clone:C430042D23 product:HYPOTHETICAL 103.9 kDa PROTEIN
DB homolog (Fragment).
GN Name=AU040320;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

Db	743	DLAVVKYLWTRDEHSLAAGVIVADTDKBPVMLLTNLVQGRYVFTLTIVSDDQGLTSSDTVS	802
Qy	806	VEVQDPPKSGLVELTLQGVGQLTEQRKDTLVRQLAVLLNVLDSDIKVKQKIR--AHSDL	863
Db	803	VNVRDPKLLNLVQMTLPMGISVLVQSELDVYVQKLQLLGD-ENKIQVRELKYDLHTD-	860
Qy	864	STVIVFYVQSRPPFKVLKAAEVARNLHMLSLSKKADFLLFKVLRVDTAGCLLKCSGHGHC	923
Db	861	ATVLVFYNDGGG-KALDGLQVERQLRTLOLQXDASILGAPAV-DIRTSVCQSDCSGHGSC	918
Qy	924	DPLTKRCICSHLMENLIQRYIWDG-ESNCWMSIFYVTVLAFTLIVLTGGFTWLCICCCCK	982
Db	919	NPITRACICEAFWMP--AGYFFNNQEA NCWSILYVFGVIVGCLLLSGVFWGIACACR	976
Qy	983	ROKRTKIRKKT-KYTILDNMDEQERMELRPKYGIKHRSTEHNSSLMVSESEFSDSDTIF	1041
Db	977	QSKKPLRQKVKYSLIGNKDEAA-----NYSRNTSLTESETDSDVLF	1020
Qy	1042	SREKMERGNPKVSMNGSIRNG	1062
Db	1021	ETRTKSNGLGKHKSHNSHSG	1041

Search completed: October 12, 2005, 10:21:38
Job time : 134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:00:24 ; Search time 908 Seconds
(without alignments)
491.154 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 5580

Sequence: 1 MAPPTGVLSLLLVITAGC.....VSMNGSIRNGASFSYCSKDR 1072

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1854112 seqs, 416015017 residues

Total number of hits satisfying chosen parameters: 1854112

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5576	99.9	1072	16 US-10-408-765A-1622	Sequence 1622, Ap
2	5576	99.9	1072	16 US-10-764-390-11	Sequence 11, Appl
3	5574	99.9	1072	16 US-10-764-390-3	Sequence 3, Appli
4	5574	99.9	1072	16 US-10-764-390-5	Sequence 5, Appli
5	5574	99.9	1072	16 US-10-764-390-8	Sequence 8, Appli
6	5574	99.9	1072	16 US-10-764-390-9	Sequence 9, Appli
7	5574	99.9	1072	16 US-10-764-390-12	Sequence 12, Appl
8	5574	99.9	1072	16 US-10-764-390-259	Sequence 259, App
9	5485	98.3	1053	16 US-10-764-390-273	Sequence 273, App
10	5485	98.3	1053	16 US-10-764-390-274	Sequence 274, App
11	5485	98.3	1063	16 US-10-764-390-7	Sequence 7, Appli

12	5485	98.3	1063	16 US-10-764-390-10	Sequence 10, Appl
13	5485	98.3	1063	16 US-10-764-390-272	Sequence 272, App
14	1451.5	26.0	1069	20 US-11-097-143-16884	Sequence 16884, A
15	1375	24.6	711	18 US-10-450-763-39353	Sequence 39353, A
16	272	4.9	155	14 US-10-050-704-175	Sequence 175, App
17	272	4.9	155	16 US-10-798-512-175	Sequence 175, App
18	232	4.2	1204	15 US-10-282-122A-49627	Sequence 49627, A
19	225	4.0	27118	17 US-10-805-684-110	Sequence 110, App
20	224	4.0	26926	9 US-09-759-508B-2	Sequence 2, Appli
21	224	4.0	26926	16 US-10-723-860-133	Sequence 133, App
22	224	4.0	26926	16 US-10-656-873A-2	Sequence 2, Appli
23	224	4.0	26926	18 US-10-756-149-4768	Sequence 4768, Ap
24	219.5	3.9	792	17 US-10-875-518-11	Sequence 11, Appl
25	216	3.9	2402	16 US-10-661-809-20	Sequence 20, Appl
26	214	3.8	36946	18 US-10-840-512-155	Sequence 155, App
27	213	3.8	7107	20 US-11-097-143-1224	Sequence 1224, Ap
28	209.5	3.8	5559	15 US-10-282-122A-75177	Sequence 75177, A
29	199	3.6	2768	20 US-11-097-143-31983	Sequence 31983, A
30	199	3.6	5701	9 US-09-864-761-37319	Sequence 37319, A
31	199	3.6	5701	14 US-10-029-386-32438	Sequence 32438, A
32	197	3.5	361	10 US-09-989-442-126	Sequence 126, App
33	197	3.5	447	17 US-10-489-695-21	Sequence 21, Appl
34	197	3.5	468	16 US-10-470-390A-18	Sequence 18, Appl
35	197	3.5	468	17 US-10-489-695-22	Sequence 22, Appl
36	197	3.5	500	15 US-10-203-875-4	Sequence 4, Appli
37	197	3.5	500	15 US-10-297-639-5	Sequence 5, Appli
38	197	3.5	550	15 US-10-369-493-10236	Sequence 10236, A
39	195	3.5	888	18 US-10-450-763-44262	Sequence 44262, A
40	195	3.5	888	18 US-10-450-763-54269	Sequence 54269, A
41	191.5	3.4	2803	15 US-10-415-187-5	Sequence 5, Appli
42	191	3.4	1194	15 US-10-282-122A-46163	Sequence 46163, A
43	188.5	3.4	1194	15 US-10-282-122A-46577	Sequence 46577, A
44	188	3.4	528	10 US-09-840-746-20	Sequence 20, Appl
45	187.5	3.4	3507	15 US-10-369-493-5784	Sequence 5784, Ap

ALIGNMENTS

RESULT 1

US-10-408-765A-1622
; Sequence 1622, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1622
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-1622

Query Match 99.9%; Score 5576; DB 16; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MAPPTGVLSLLLVITAGCARKQCSGRYTSNAVISPNLETTTRIMRVSHYTFPVVDCTAA 60
Db 1 MAPPTGVLSLLLVITAGCARKQCSGRYTSNAVISPNLETTTRIMRVSHYTFPVVDCTAA 60
Qy 61 CCDSLCDLAWFEGRCYLVSCPHKENCEPKMGPIRSYLTFLVRPVQRPQALLDYGDM 120

Db	61	CCDUSCDLAWFEGRCYLVSCPHKENCEPKMGPIRISYLTFTVLURPVORPAQLLDYDGM	120
Qy	121	LNRGSPGIGWSDSPEDIRKDLXFLGDKWGLEEMSEYXDDYREKOLLQPSGKQEPGRGA	180
Db	121	LNRGSPGIGWSDSPEDIRKDLTFLGDKWGLEEMSEYSDDYREKOLLQPSGKQEPGRGA	180
Qy	181	EYTDWGLLPGSEGFANFSGVSDSPAVPAETQODPLHLYNLSASTPAPKLPERSVLLPLPT	240
Db	181	EYTDWGLLPGSEGFANFSGVSDSPAVPAETQODPELHLYNLSASTPAPKLPERSVLLPLPT	240
Qy	241	TPSSEVLEKEKASQLOEQSSNGSKEVLPWSPHSLPPASLELSVTVEKSPVLTVTPGST	300
Db	241	TPSSEVLEKEKASQLOEQSSNGSKEVLPWSPHSLPPASLELSVTVEKSPVLTVTPGST	300
Qy	301	EHSIPTPTTSAAPSESTPSELPIPTTAPTARTVKELTVSAGDNLITLIPDNEVELKAFVAP	360
Db	301	EHSIPTPTTSAAPSESTPSELPIPTTAPTARTVKELTVSAGDNLITLIPDNEVELKAFVAP	360
Qy	361	APPVETTYNYEWNLLSHPTDQGEIKOGHKOTLNLSQLSVGLYVFKVTVSENAFGGFV	420
Db	361	APPVETTYNYEWNLLSHPTDQGEIKOGHKOTLNLSQLSVGLYVFKVTVSENAFGGFV	420
Qy	421	NVTVKPARRVNLPPVAVVVSQLOELTLPLTSALIDGQSTDDETVISVHWEENGPPIEE	480
Db	421	NVTVKPARRVNLPPVAVVVSQLOELTLPLTSALIDGQSTDDETVISVHWEENGPPIEE	480
Qy	481	KTSVDSVPLRLSNLDPGNYSFRLTVTDSDGATNSTTAALIVNNAVDPYPVANAGPNHTIT	540
Db	481	KTSVDSVPLRLSNLDPGNYSFRLTVTDSDGATNSTTAALIVNNAVDPYPVANAGPNHTIT	540
Qy	541	LPQNSITILNGQSSDDHQIVLYEWSLPGSGEKHVMQGVQTPYLHLASMQEGDYTEQLK	600
Db	541	LPQNSITILNGQSSDDHQIVLYEWSLPGSGEKHVMQGVQTPYLHLASMQEGDYTEQLK	600
Qy	601	VTDSSRQOSTAXTVTVIQQENRPPVAVAGDPKELIPPVESATLDGSSSDDHGIVPYHW	660
Db	601	VTDSSRQOSTAVTVTVIQQENRPPVAVAGDPKELIPPVESATLDGSSSDDHGIVPYHW	660
Qy	661	EHVRGPAVEMENIDKAIATVTGLOVGYTHFRLTVKQOGLSSTSTLTAVAKKENNSPPR	720
Db	661	EHVRGPAVEMENIDKAIATVTGLOVGYTHFRLTVKQOGLSSTSTLTAVAKKENNSPPR	720
Qy	721	ARAGRHVLVLPNNNSITLDGSRSTDQRIVSYLWIRDGQSPAAGDVIDGSDHSVALQTN	780
Db	721	ARAGRHVLVLPNNNSITLDGSRSTDQRIVSYLWIRDGQSPAAGDVIDGSDHSVALQTN	780
Qy	781	LVEGVYTHLAVTDSQNASDTDATVEQOPPRKSGLVELTLOVGVQOLTEQRKDTLVRQ	840
Db	781	LVEGVYTHLAVTDSQNASDTDATVEQOPPRKSGLVELTLOVGVQOLTEQRKDTLVRQ	840
Qy	841	LAVLLNVLDSDIKVQIRAHSDLSITVIFYVQSRPPPKVLKAAEVARNLHWRLSKEKADF	900
Db	841	LAVLLNVLDSDIKVQIRAHSDLSITVIFYVQSRPPPKVLKAAEVARNLHWRLSKEKADF	900
Qy	901	LLFKVLRVDTAGCLLKCSGHGCHDPLTKRCITCSHLWMENLIRQYIWDGSENCEWSIIFYVT	960
Db	901	LLFKVLRVDTAGCLLKCSGHGCHDPLTKRCITCSHLWMENLIRQYIWDGSENCEWSIIFYVT	960
Qy	961	VLAFTLIVLTGFTWLCTCCCKROKRTKIRKTKYTLTDNMDQBERMELRPKYGIKIRST	1020
Db	961	VLAFTLIVLTGFTWLCTCCCKROKRTKIRKTKYTLTDNMDQBERMELRPKYGIKIRST	1020
Qy	1021	EHNSSLMWSESEFSDQDTTIFSRKMERGNPKVSMNGSIRNGASFYSYCSKDR	1072
Db	1021	EHNSSLMWSESEFSDQDTTIFSRKMERGNPKVSMNGSIRNGASFYSYCSKDR	1072

```

; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-11

Query Match          99.9%; Score 5576; DB 16; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps

```

Db 661 EHVGRPSAVEMENIDKAIATVGLQVGYHFLRVKQOGLSSSTLTAVKKNNSPPR 720
Qy 721 ARAGGRHVLVLPNNISITLDGSRSDQRIYVSYLWIRGQSPAGDVIDGSDHVALQLTN 780
Db 721 ARAGGRHVLVLPNNISITLDGSRSDQRIYVSYLWIRGQSPAGDVIDGSDHVALQLTN 780
Qy 781 LVEGVYTHLRVTDQASDQDTATVEVQDPKRSGLVELTLQVGVQLTQKDTLVRQ 840
Db 781 LVEGVYTHLRVTDQASDQDTATVEVQDPKRSGLVELTLQVGVQLTQKDTLVRQ 840
Qy 841 LAVLLNVLDSDIKVQKIRAHSDLSSTVIVFYVQSRPPPKVLKAAEVARNLHRLSKKADF 900
Db 841 LAVLLNVLDSDIKVQKIRAHSDLSSTVIVFYVQSRPPPKVLKAAEVARNLHRLSKKADF 900
Qy 901 LLFKVLVDTAGCLLKCSGHGCDPLTKRCICSHLWMENLIQRYWIDGESNCEWSIFVVT 960
Db 901 LLFKVLVDTAGCLLKCSGHGCDPLTKRCICSHLWMENLIQRYWIDGESNCEWSIFVVT 960
Qy 961 VLAFTLVLVTGGFTWLCICCKKQKRTKIRKTKYTTILDNDQERMELRPKYGIKHRST 1020
Db 961 VLAFTLVLVTGGFTWLCICCKKQKRTKIRKTKYTTILDNDQERMELRPKYGIKHRST 1020
Qy 1021 EHNSSLVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSCKDR 1072
Db 1021 EHNSSLVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSCKDR 1072

RESULT 3

US-10-764-390-3
; Sequence 3, Application US/10764390
; Publication No. US2004021421A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254P16B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 142
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 157
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 612
; OTHER INFORMATION: Xaa = any amino acid
; US-10-764-390-3

Query Match 99.9%; Score 5574; DB 16; Length 1072;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1072; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPPTGVLSSILLVTTIAGCARCOCSEGRYSNAVISPNLETTIMRVSHTPFVVDCTAA 60
Db 1 MAPPTGVLSSILLVTTIAGCARCOCSEGRYSNAVISPNLETTIMRVSHTPFVVDCTAA 60
Qy 61 CCDLSSCDLAWFEGRCYLVSCPHKENCCEPKGPIRSYLTFTVLRPVQRPQAQLLDYDMM 120
Db 61 CCDLSSCDLAWFEGRCYLVSCPHKENCCEPKGPIRSYLTFTVLRPVQRPQAQLLDYDMM 120
Qy 121 LNRGSPSGIWDGSPEDIRKDLXFLGKDWGLLEBSEYXDDYRELEKDLLQSGKQEPGSA 180
Db 121 LNRGSPSGIWDGSPEDIRKDLXFLGKDWGLLEBSEYXDDYRELEKDLLQSGKQEPGSA 180
Qy 181 EYTDWGLLPQSEGAFFNSVGDSPAVPAETQODPELHYLNEASAPKLPERSVLLPLPT 240
Db 181 EYTDWGLLPQSEGAFFNSVGDSPAVPAETQODPELHYLNEASAPKLPERSVLLPLPT 240
Qy 241 TPSSGEVLEKEKASQLQEQSSNSGKEVLMPSHSLPPASLELSVTVEKSPVLTVTGST 300
Db 241 TPSSGEVLEKEKASQLQEQSSNSGKEVLMPSHSLPPASLELSVTVEKSPVLTVTGST 300
Qy 301 EHSIPTPTTAAPESESTPSELPIPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
Db 301 EHSIPTPTTAAPESESTPSELPIPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
Qy 361 APPVETTYNENLISHPTDYQGEIKQGHKQTLNLSQSVGLYVFKVTVSSENAFGEFV 420
Db 361 APPVETTYNENLISHPTDYQGEIKQGHKQTLNLSQSVGLYVFKVTVSSENAFGEFV 420
Qy 421 NVTVKPARRVNLPPVAVVSPQLQELTLPLTSALIDGSDTDDTEIVSYHMEENGPTEE 480
Db 421 NVTVKPARRVNLPPVAVVSPQLQELTLPLTSALIDGSDTDDTEIVSYHMEENGPTEE 480
Qy 481 KTSVDSVPLRLSLNDPGNYSFRLTVTSDGATNSTTAALIVNNAVDYPPVANAGNHTIT 540
Db 481 KTSVDSVPLRLSLNDPGNYSFRLTVTSDGATNSTTAALIVNNAVDYPPVANAGNHTIT 540
Qy 541 LPQNSITLNGQSDDDHQIVLYEWSLPGSGEGKHVVMQGVQTPVYLHLSAMQEGDYTFOLK 600
Db 541 LPQNSITLNGQSDDDHQIVLYEWSLPGSGEGKHVVMQGVQTPVYLHLSAMQEGDYTFOLK 600
Qy 601 VTDSSROQSTAXTVIVQPNRPPVAVAGPKELIFPVESATLDGSSSDHGHGVFVHW 660
Db 601 VTDSSROQSTAXTVIVQPNRPPVAVAGPKELIFPVESATLDGSSSDHGHGVFVHW 660
Qy 661 EHVGRPSAVEMENIDKAIATVGLQVGYHFLRVKQOGLSSSTLTAVKKNNSPPR 720
Db 661 EHVGRPSAVEMENIDKAIATVGLQVGYHFLRVKQOGLSSSTLTAVKKNNSPPR 720
Qy 721 ARAGGRHVLVLPNNISITLDGSRSDQRIYVSYLWIRGQSPAGDVIDGSDHVALQLTN 780
Db 721 ARAGGRHVLVLPNNISITLDGSRSDQRIYVSYLWIRGQSPAGDVIDGSDHVALQLTN 780
Qy 781 LVEGVYTHLRVTDQASDQDTATVEVQDPKRSGLVELTLQVGVQLTQKDTLVRQ 840
Db 781 LVEGVYTHLRVTDQASDQDTATVEVQDPKRSGLVELTLQVGVQLTQKDTLVRQ 840
Qy 841 LAVLLNVLDSDIKVQKIRAHSDLSSTVIVFYVQSRPPPKVLKAAEVARNLHRLSKKADF 900
Db 841 LAVLLNVLDSDIKVQKIRAHSDLSSTVIVFYVQSRPPPKVLKAAEVARNLHRLSKKADF 900
Qy 901 LLFKVLVDTAGCLLKCSGHGCDPLTKRCICSHLWMENLIQRYWIDGESNCEWSIFVVT 960
Db 901 LLFKVLVDTAGCLLKCSGHGCDPLTKRCICSHLWMENLIQRYWIDGESNCEWSIFVVT 960
Qy 961 VLAFTLVLVTGGFTWLCICCKKQKRTKIRKTKYTTILDNDQERMELRPKYGIKHRST 1020
Db 961 VLAFTLVLVTGGFTWLCICCKKQKRTKIRKTKYTTILDNDQERMELRPKYGIKHRST 1020
Qy 1021 EHNSSLVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSCKDR 1072
Db 1021 EHNSSLVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFYSCKDR 1072

```

RESULT 4
US-10-764-390-5
; Sequence 5, Application US/10764390
; Publication No. US2004021412AI
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-5

Query Match          99.9%; Score 5574; DB 16; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPPTGVLSSLLLVTIAGCARKQCSEGRYTSNAVISPNELETTRIMRVSHTEPPVVDCTAA 60
DB 1 MAPPTGVLSSLLLVTIAGCARKQCSEGRYTSNAVISPNELETTRIMRVSHTEPPVVDCTAA 60
QY 61 CDDLSDDLAWFEGRCYLVSCEPHKENCPEKMGPIRSYLTFLVPVQRPALQDYGDM 120
DB 61 CDDLSDDLAWFEGRCYLVSCEPHKENCPEKMGPIRSYLTFLVPVQRPALQDYGDM 120
QY 121 LNRGSPSGIWDGSPEDIRKDLFLGKDWGLEMSYDDYRELEKDLQPSGKQPRGSA 180
DB 121 LNRGSPSGIWDGSPEDIRKDLFLGKDWGLEMSYDDYRELEKDLQPSGKQPRGSA 180
QY 181 EYTDWGLLPGSGAFNSVGDSPAVPAETQDPPELHYLNESASTAPKLPERSVLLPPT 240
DB 181 EYTDWGLLPGSGAFNSVGDSPAVPAETQDPPELHYLNESASTAPKLPERSVLLPPT 240
QY 241 TPSSGEVLEKEKASQLOEQSSNGKEVLMPSHSLPPASLELSSVTVKESPVLTTPGST 300
DB 241 TPSSGEVLEKEKASQLOEQSSNGKEVLMPSHSLPPASLELSSVTVKESPVLTTPGST 300
QY 301 EHSIPTPTSAAPSESTPSELPSPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
DB 301 EHSIPTPTSAAPSESTPSELPSPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
QY 361 APPVETTYNENLISHPTDYQGEIKQGHKOTLINLSQSLVGLYVFKVTVSSNAFGEFV 420
DB 361 APPVETTYNENLISHPTDYQGEIKQGHKOTLINLSQSLVGLYVFKVTVSSNAFGEFV 420
QY 421 NVTVPARRVNLPPVAVVSPQLQELTLPLTSALIDGQSQTDDTEIVSYHWEENPGFTIEE 480
DB 421 NVTVPARRVNLPPVAVVSPQLQELTLPLTSALIDGQSQTDDTEIVSYHWEENPGFTIEE 480
QY 481 KTSVDSVPLRLSNLDGNSYFRLTVTSDGATNSTTAAIIVNNAVDPYPPVANAGNHTIT 540
DB 481 KTSVDSVPLRLSNLDGNSYFRLTVTSDGATNSTTAAIIVNNAVDPYPPVANAGNHTIT 540
QY 541 LPQNSITLNGQSSDDHQIIVLWESLPGSGEGKHVVMQGVQTPYLHLSAMQSGDITFQJK 600
DB 541 LPQNSITLNGQSSDDHQIIVLWESLPGSGEGKHVVMQGVQTPYLHLSAMQSGDITFQJK 600

RESULT 5
US-10-764-390-8
; Sequence 8, Application US/10764390
; Publication No. US2004021412AI
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-8

Query Match          99.9%; Score 5574; DB 16; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPPTGVLSSLLLVTIAGCARKQCSEGRYTSNAVISPNELETTRIMRVSHTEPPVVDCTAA 60
DB 1 MAPPTGVLSSLLLVTIAGCARKQCSEGRYTSNAVISPNELETTRIMRVSHTEPPVVDCTAA 60
QY 61 CDDLSDDLAWFEGRCYLVSCEPHKENCPEKMGPIRSYLTFLVPVQRPALQDYGDM 120
DB 61 CDDLSDDLAWFEGRCYLVSCEPHKENCPEKMGPIRSYLTFLVPVQRPALQDYGDM 120
QY 121 LNRGSPSGIWDGSPEDIRKDLFLGKDWGLEMSYDDYRELEKDLQPSGKQPRGSA 180
DB 121 LNRGSPSGIWDGSPEDIRKDLFLGKDWGLEMSYDDYRELEKDLQPSGKQPRGSA 180
QY 181 EYTDWGLLPGSGAFNSVGDSPAVPAETQDPPELHYLNESASTAPKLPERSVLLPPT 240
DB 181 EYTDWGLLPGSGAFNSVGDSPAVPAETQDPPELHYLNESASTAPKLPERSVLLPPT 240
QY 241 TPSSGEVLEKEKASQLOEQSSNGKEVLMPSHSLPPASLELSSVTVKESPVLTTPGST 300
DB 241 TPSSGEVLEKEKASQLOEQSSNGKEVLMPSHSLPPASLELSSVTVKESPVLTTPGST 300
QY 301 EHSIPTPTSAAPSESTPSELPSPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
DB 301 EHSIPTPTSAAPSESTPSELPSPTTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
QY 361 APPVETTYNENLISHPTDYQGEIKQGHKOTLINLSQSLVGLYVFKVTVSSNAFGEFV 420
DB 361 APPVETTYNENLISHPTDYQGEIKQGHKOTLINLSQSLVGLYVFKVTVSSNAFGEFV 420
QY 421 NVTVPARRVNLPPVAVVSPQLQELTLPLTSALIDGQSQTDDTEIVSYHWEENPGFTIEE 480
DB 421 NVTVPARRVNLPPVAVVSPQLQELTLPLTSALIDGQSQTDDTEIVSYHWEENPGFTIEE 480
QY 481 KTSVDSVPLRLSNLDGNSYFRLTVTSDGATNSTTAAIIVNNAVDPYPPVANAGNHTIT 540
DB 481 KTSVDSVPLRLSNLDGNSYFRLTVTSDGATNSTTAAIIVNNAVDPYPPVANAGNHTIT 540
QY 541 LPQNSITLNGQSSDDHQIIVLWESLPGSGEGKHVVMQGVQTPYLHLSAMQSGDITFQJK 600
DB 541 LPQNSITLNGQSSDDHQIIVLWESLPGSGEGKHVVMQGVQTPYLHLSAMQSGDITFQJK 600

601 VTDSSRQOSTAXVTVIQVQENNRPPVAVAGDPKELIFPVESATLDGSSSSDDHGIVFYHW 660
601 VTDSSRQOSTAXVTVIQVQENNRPPVAVAGDPKELIFPVESATLDGSSSSDDHGIVFYHW 660
661 EHVGPSAVEMENIDKAIATVTGLQVGYTHFRLTVKQOQGLSSSTLTAVAKKENNSPPR 720
661 EHVGPSAVEMENIDKAIATVTGLQVGYTHFRLTVKQOQGLSSSTLTAVAKKENNSPPR 720
721 ARAGRHVLVLPNNSITLDGSRSTDORIVSYLWIRDOQSPAAGDVIDGSDHSHVALQLTN 780
721 ARAGRHVLVLPNNSITLDGSRSTDORIVSYLWIRDOQSPAAGDVIDGSDHSHVALQLTN 780
781 LVEGVYTFHLRVTDQSGASDTDTATVEVQDPKRSGLVELTLQVGVGLTEQRKDTLVRQ 840
781 LVEGVYTFHLRVTDQSGASDTDTATVEVQDPKRSGLVELTLQVGVGLTEQRKDTLVRQ 840
841 LAVLLNVLDSDIKVQKIRAHSDLSSTVIIVFYQSPPPFKVLKAAEVARNLHRLSKEKADF 900
841 LAVLLNVLDSDIKVQKIRAHSDLSSTVIIVFYQSPPPFKVLKAAEVARNLHRLSKEKADF 900
901 LLFKVLRVDTAGCLLKSGHGHCDPLTKRCICSHLWMENLIQRYIWDGESNCESIFVYT 960
901 LLFKVLRVDTAGCLLKSGHGHCDPLTKRCICSHLWMENLIQRYIWDGESNCESIFVYT 960
961 VLAFTLIVLTGFTWLCICCCCKRQKRTKIRKTKYITLDNMDEQERMELRPKYGIKIRST 1020
961 VLAFTLIVLTGFTWLCICCCCKRQKRTKIRKTKYITLDNMDEQERMELRPKYGIKIRST 1020
1021 EHNSLWSESEFSDQDTIFSRKMERGNPKVMNGSIRNGASFYSCKDR 1072
1021 EHNSLWSESEFSDQDTIFSRKMERGNPKVMNGSIRNGASFYSCKDR 1072

```

```
Qy 121 LNRGSPGSGWGDSPEDIRKDLXFLGKDWGLGLEEMSEYXDDYRELEKDLLOPSGKQEPGRSA 180
Db 121 LNRGSPGSGWGDSPEDIRKDLXFLGKDWGLGLEEMSEYXDDYRELEKDLLOPSGKQEPGRSA 180
Qy 181 EYTDWGLLPGSEGAFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
Db 181 EYTDWGLLPGSEGAFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
Qy 241 TPSSGEVLEKASQLOEQSSNSGKEVLPSPHSLPPASLELSVTVVEKSPVLVTPGST 300
Db 241 TPSSGEVLEKASQLOEQSSNSGKEVLPSPHSLPPASLELSVTVVEKSPVLVTPGST 300
Qy 301 EHSIPTPTSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLITLTPDNEVELKAFVAP 360
Db 301 EHSIPTPTSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLITLTPDNEVELKAFVAP 360
Qy 361 APPVETTYNWEWNLISHPTDYQGEIKQGHKQTLNLSQSLVGLYVFKVTVSSENAFGEGFV 420
Db 361 APPVETTYNWEWNLISHPTDYQGEIKQGHKQTLNLSQSLVGLYVFKVTVSSENAFGEGFV 420
Qy 421 NVTVKPARRVNLPPVAVVSPOLQBELTLPSTAPRTVKELTVSAGDNLITLTPDNEVELKAFVAP 480
Db 421 NVTVKPARRVNLPPVAVVSPOLQBELTLPSTAPRTVKELTVSAGDNLITLTPDNEVELKAFVAP 480
Qy 481 KTSVDSVPLRLSNDLPCGNYSFRLTVTSDGATNSTTAAALIVNNAVDPYPVANAGPNHTIT 540
Db 481 KTSVDSVPLRLSNDLPCGNYSFRLTVTSDGATNSTTAAALIVNNAVDPYPVANAGPNHTIT 540
Qy 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGKHVMQGVQVTPYLHLHSAMQEGDYTFQLK 600
Db 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGKHVMQGVQVTPYLHLHSAMQEGDYTFQLK 600
Qy 601 VTDSSROQSTAXVTVI VQENNRPPVAVAGDPKELIPPVESATLDGSSSSDDHGIVFYHW 660
Db 601 VTDSSROQSTAXVTVI VQENNRPPVAVAGDPKELIPPVESATLDGSSSSDDHGIVFYHW 660
Qy 721 ARAGGRHVLPNNISITLDGSRSDTDORI VSYLWIRDGQSPAAGDVIDGSDHVALQLTN 780
Db 721 ARAGGRHVLPNNISITLDGSRSDTDORI VSYLWIRDGQSPAAGDVIDGSDHVALQLTN 780
Qy 781 LVEGVYTFHLRVTD SQASDSDTDTATVEVQDPDKSGLVLTQVGVGQLTQQRKDTLVRQ 840
Db 781 LVEGVYTFHLRVTD SQASDSDTDTATVEVQDPDKSGLVLTQVGVGQLTQQRKDTLVRQ 840
Qy 841 LAVLLNLVDSDIKVQKTRAHSDLSVTIVVYVQSRPPPKVLKAAEVARNLHMLRSLKEKADF 900
Db 841 LAVLLNLVDSDIKVQKTRAHSDLSVTIVVYVQSRPPPKVLKAAEVARNLHMLRSLKEKADF 900
Qy 901 LLEFKVLVDTAGCLLKCSGHGCHDPLTKRCICSHLWENLQRYIWDGESNCWSIFVVT 960
Db 901 LLEFKVLVDTAGCLLKCSGHGCHDPLTKRCICSHLWENLQRYIWDGESNCWSIFVVT 960
Qy 961 VLAFTLIVLTGGFTWLCICCCCKRQKRTKIRKTKYTTILDNMDEQERMLRKYGIKHRST 1020
Db 961 VLAFTLIVLTGGFTWLCICCCCKRQKRTKIRKTKYTTILDNMDEQERMLRKYGIKHRST 1020
Qy 1021 EHNSLWVSESEFSDQDTIFSRKMERGNPKVSNMGSIRNGASFYSCKDR 1072
Db 1021 EHNSLWVSESEFSDQDTIFSRKMERGNPKVSNMGSIRNGASFYSCKDR 1072
```

RESULT 6

```
US-10-764-390-9
; Sequence 9, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
```

```
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-9
```

```
Query Match 99.9%; Score 5574; DB 16; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 MAPPTGVLSSLLLVTTIAGCARKQCSGRTYSNAVISPNLETTIRIMRVSHTFPVDCTAA 60
Db 1 MAPPTGVLSSLLLVTTIAGCARKQCSGRTYSNAVISPNLETTIRIMRVSHTFPVDCTAA 60
Qy 61 CCDSLSSCDLAWPFEGRCVLSVCPHKENCEPKMGPIRSYLTFLVRPQORPAQLLDYGDMM 120
Db 61 CCDSLSSCDLAWPFEGRCVLSVCPHKENCEPKMGPIRSYLTFLVRPQORPAQLLDYGDMM 120
Qy 121 LNRGSPGSGWGDSPEDIRKDLXFLGKDWGLGLEEMSEYXDDYRELEKDLLOPSGKQEPGRSA 180
Db 121 LNRGSPGSGWGDSPEDIRKDLXFLGKDWGLGLEEMSEYXDDYRELEKDLLOPSGKQEPGRSA 180
Qy 181 EYTDWGLLPGSEGAFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
Db 181 EYTDWGLLPGSEGAFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
Qy 241 TPSSGEVLEKASQLOEQSSNSGKEVLPSPHSLPPASLELSVTVVEKSPVLVTPGST 300
Db 241 TPSSGEVLEKASQLOEQSSNSGKEVLPSPHSLPPASLELSVTVVEKSPVLVTPGST 300
Qy 301 EHSIPTPTSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLITLTPDNEVELKAFVAP 360
Db 301 EHSIPTPTSAAPSESTPSELPISTPTAPRTVKELTVSAGDNLITLTPDNEVELKAFVAP 360
Qy 361 APPVETTYNWEWNLISHPTDYQGEIKQGHKQTLNLSQSLVGLYVFKVTVSSENAFGEGFV 420
Db 361 APPVETTYNWEWNLISHPTDYQGEIKQGHKQTLNLSQSLVGLYVFKVTVSSENAFGEGFV 420
Qy 421 NVTVKPARRVNLPPVAVVSPOLQBELTLPSTAPRTVKELTVSAGDNLITLTPDNEVELKAFVAP 480
Db 421 NVTVKPARRVNLPPVAVVSPOLQBELTLPSTAPRTVKELTVSAGDNLITLTPDNEVELKAFVAP 480
Qy 481 KTSVDSVPLRLSNDLPCGNYSFRLTVTSDGATNSTTAAALIVNNAVDPYPVANAGPNHTIT 540
Db 481 KTSVDSVPLRLSNDLPCGNYSFRLTVTSDGATNSTTAAALIVNNAVDPYPVANAGPNHTIT 540
Qy 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGKHVMQGVQVTPYLHLHSAMQEGDYTFQLK 600
Db 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGKHVMQGVQVTPYLHLHSAMQEGDYTFQLK 600
Qy 601 VTDSSROQSTAXVTVI VQENNRPPVAVAGDPKELIPPVESATLDGSSSSDDHGIVFYHW 660
Db 601 VTDSSROQSTAXVTVI VQENNRPPVAVAGDPKELIPPVESATLDGSSSSDDHGIVFYHW 660
Qy 661 EHVGRPSAVEMENIDKAIATVTGLQVGTYHFLRTVKDQOGLSSTSTLTVA VKKENNSPPR 720
Db 661 EHVGRPSAVEMENIDKAIATVTGLQVGTYHFLRTVKDQOGLSSTSTLTVA VKKENNSPPR 720
```

```

QY 721 ARAGGRHVLVLPNNSITLDGSRSTDDQRIYSYLWIRDGQSPAAGDVIDGSDHSHVALQLTN 780
Db 721 ARAGGRHVLVLPNNSITLDGSRSTDDQRIYSYLWIRDGQSPAAGDVIDGSDHSHVALQLTN 780
QY 781 LVEGVYTHLRLVTDGQASDQDTATVEVQDPKRSGLVELTLQVGVGQLTQQRKDTLVRQ 840
Db 781 LVEGVYTHLRLVTDGQASDQDTATVEVQDPKRSGLVELTLQVGVGQLTQQRKDTLVRQ 840
QY 841 LAVLLNLVLDSDIKVQKIRAHSDLSVTIVFYVQSRPPFKVLAEEVARNLHMRLSKEKADF 900
Db 841 LAVLLNLVLDSDIKVQKIRAHSDLSVTIVFYVQSRPPFKVLAEEVARNLHMRLSKEKADF 900
QY 901 LFFKVLRLVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCESIFVYT 960
Db 901 LFFKVLRLVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCESIFVYT 960
QY 961 VLAFTLIVLTGGFTWLCICCCCKRQRTKIRKTKYTILDNMDEQRMELRPKYGIKHRST 1020
Db 961 VLAFTLIVLTGGFTWLCICCCCKRQRTKIRKTKYTILDNMDEQRMELRPKYGIKHRST 1020
QY 1021 EHNSSLMVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFSYCSKOR 1072
Db 1021 EHNSSLMVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFSYCSKOR 1072

RESULT 7
US-10-764-390-12
; Sequence 12, Application US/10764390
; Publication No. US2004021421A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254p1D6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1072
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-764-390-12

Query Match 99.9%; Score 5574; DB 16; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPPTGVLSLLLVITAGCARQCSGRTYNAVISPNELETRIMRVSHHTPPVVDCTAA 60
Db 1 MAPPTGVLSLLLVITAGCARQCSGRTYNAVISPNELETRIMRVSHHTPPVVDCTAA 60
QY 61 CDDLSLCLAWPEGRCYLVSCPHKENCPEKMGPIRSYLTFLVRVQRPQALLDYGDMM 120
Db 61 CDDLSLCLAWPEGRCYLVSCPHKENCPEKMGPIRSYLTFLVRVQRPQALLDYGDMM 120
QY 121 LNRGSPGIWGSPEIRKDLPLGKDWGLEEMSEYKDDYRELEKDLQPSGKQPRGSA 180
Db 121 LNRGSPGIWGSPEIRKDLPLGKDWGLEEMSEYKDDYRELEKDLQPSGKQPRGSA 180
QY 181 EYTDWGLLPGSEGAFFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240

```

```

Db 181 EYTDWGLLPGSEGAFFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPT 240
QY 241 TPSSGEVLEKEKASQLOEQSSNSGKEVLMPSHSLPPASLELSVTVYKESVLTVPGST 300
Db 241 TPSSGEVLEKEKASQLOEQSSNSGKEVLMPSHSLPPASLELSVTVYKESVLTVPGST 300
QY 301 EHSIPTPTGAAAPSESTPSELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
Db 301 EHSIPTPTGAAAPSESTPSELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAP 360
QY 361 APPVETTYNTEWNLISHTDYQGBIKQGHKOTLNLSQSVGLYVFKVTVSSENAFGEFV 420
Db 361 APPVETTYNTEWNLISHTDYQGBIKQGHKOTLNLSQSVGLYVFKVTVSSENAFGEFV 420
QY 421 NVTVKPARRVNLPPVAVVSPQLQELTLPLTSALIDGSDQSTDDTEIVSYHWEIINGPFTIEE 480
Db 421 NVTVKPARRVNLPPVAVVSPQLQELTLPLTSALIDGSDQSTDDTEIVSYHWEIINGPFTIEE 480
QY 481 KTSVDSVPLRLSNLDPNGYSFRLTVTSDGATNSTTAAALIVNNAVDPYPPVANAGPNHTIT 540
Db 481 KTSVDSVPLRLSNLDPNGYSFRLTVTSDGATNSTTAAALIVNNAVDPYPPVANAGPNHTIT 540
QY 541 LPQNSITLNGQSSDDHQIIVLYEWSLGPSEKGVVMQGVQTPYLHLSAMQEGDITFOLK 600
Db 541 LPQNSITLNGQSSDDHQIIVLYEWSLGPSEKGVVMQGVQTPYLHLSAMQEGDITFOLK 600
QY 601 VTSSRQOSTAXVTVIIVQPENNRPPVAVAGDPKELIPVESATLDGSSSDHGHIVFYHW 660
Db 601 VTSSRQOSTAXVTVIIVQPENNRPPVAVAGDPKELIPVESATLDGSSSDHGHIVFYHW 660
QY 661 EHVGPSAVEMENIDKAIATVTGLQVGYTHFRLTVKQOQGLSSTSTLTVAVKENNSPPR 720
Db 661 EHVGPSAVEMENIDKAIATVTGLQVGYTHFRLTVKQOQGLSSTSTLTVAVKENNSPPR 720
QY 721 ARAGGRHVLVLPNNSITLDGSRSTDDQRIYSYLWIRDGQSPAAGDVIDGSDHSHVALQLTN 780
Db 721 ARAGGRHVLVLPNNSITLDGSRSTDDQRIYSYLWIRDGQSPAAGDVIDGSDHSHVALQLTN 780
QY 781 LVEGVYTHLRLVTDGQASDQDTATVEVQDPKRSGLVELTLQVGVGQLTQQRKDTLVRQ 840
Db 781 LVEGVYTHLRLVTDGQASDQDTATVEVQDPKRSGLVELTLQVGVGQLTQQRKDTLVRQ 840
QY 841 LAVLLNLVLDSDIKVQKIRAHSDLSVTIVFYVQSRPPFKVLAEEVARNLHMRLSKEKADF 900
Db 841 LAVLLNLVLDSDIKVQKIRAHSDLSVTIVFYVQSRPPFKVLAEEVARNLHMRLSKEKADF 900
QY 901 LFFKVLRLVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCESIFVYT 960
Db 901 LFFKVLRLVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRYIWDGESNCESIFVYT 960
QY 961 VLAFTLIVLTGGFTWLCICCCCKRQRTKIRKTKYTILDNMDEQRMELRPKYGIKHRST 1020
Db 961 VLAFTLIVLTGGFTWLCICCCCKRQRTKIRKTKYTILDNMDEQRMELRPKYGIKHRST 1020
QY 1021 EHNSSLMVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFSYCSKOR 1072
Db 1021 EHNSSLMVSESEFSDQDTIFSRKMERGNPKVSMNGSIRNGASFSYCSKOR 1072

RESULT 8
US-10-764-390-259
; Sequence 259, Application US/10764390
; Publication No. US2004021421A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins

```

; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 259
; LENGTH: 1072
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-259

Query Match 99.9%; Score 5574; DB 16; Length 1072;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1069; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPPTGVLSLLLLVTTAGCARKQCSGRTYSNAVISPNLETTIRIMRVSHTPFVVDCTAA 60
DB 1 MAPPTGVLSLLLLVTTAGCARKQCSGRTYSNAVISPNLETTIRIMRVSHTPFVVDCTAA 60
QY 61 CCDSLSSCDLAWFEGRCVLSVCPHKENCEPKMGPIRSYLTFLVRPQRPQALLDYGMM 120
DB 61 CCDSLSSCDLAWFEGRCVLSVCPHKENCEPKMGPIRSYLTFLVRPQRPQALLDYGMM 120
QY 121 LNRGSPGIGWSDPEDIRKDLFLGKOWGLEMESEYDXYRELEKDLLQPSGKQPRGSA 180
DB 121 LNRGSPGIGWSDPEDIRKDLFLGKOWGLEMESEYDXYRELEKDLLQPSGKQPRGSA 180
QY 181 EYTWGLLPGSEGAFNSVGSVPAPVAPARTQDPDELHYLNESASTPAPKLPERSVLLPLPT 240
DB 181 EYTWGLLPGSEGAFNSVGSVPAPVAPARTQDPDELHYLNESASTPAPKLPERSVLLPLPT 240
QY 241 TPSSGEVLEKASQLOQSSNSGKVLMPSHSLPPASLELSVTVVEKSPVLVTPGST 300
DB 241 TPSSGEVLEKASQLOQSSNSGKVLMPSHSLPPASLELSVTVVEKSPVLVTPGST 300
QY 301 EHSIPTPTSAAPSESTPSELPIPTTAPRTVKELTVSAGDNLITLPDNEVELKAFVAP 360
DB 301 EHSIPTPTSAAPSESTPSELPIPTTAPRTVKELTVSAGDNLITLPDNEVELKAFVAP 360
QY 361 APPVETTYNENWLIHSHTDYQGEIKQGHKOTLNLSQLSVGLYVFKVTVSSENAFGEGFV 420
DB 361 APPVETTYNENWLIHSHTDYQGEIKQGHKOTLNLSQLSVGLYVFKVTVSSENAFGEGFV 420
QY 421 NVTVKPARVNLPPVAVVSPQLBLTLPLTSALIDGSGSTDDTETIVSYHWEINGPFTIE 480
DB 421 NVTVKPARVNLPPVAVVSPQLBLTLPLTSALIDGSGSTDDTETIVSYHWEINGPFTIE 480
QY 481 KTSVDSPLRLSLNDPKNYSFRLTVTSDGATNSTTAALIVNNAVDYPPVANAGPNHTIT 540
DB 481 KTSVDSPLRLSLNDPKNYSFRLTVTSDGATNSTTAALIVNNAVDYPPVANAGPNHTIT 540
QY 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGKHVVMQGVQTPYLHLSAMQEGDYTFQLK 600
DB 541 LPQNSITLNGQSSDDHQIVLYEWSLPGSGKHVVMQGVQTPYLHLSAMQEGDYTFQLK 600
QY 601 VTDSSRQOSTAXVTIVQENNRPPVAVAGDPKELIPVESATLDGSSSSDDHGIVFVHW 660
DB 601 VTDSSRQOSTAXVTIVQENNRPPVAVAGDPKELIPVESATLDGSSSSDDHGIVFVHW 660
QY 661 EHVGPASAVENIDKALATVTLQVGYTHFLTVKDOQGLSSTSTLTVAVKKENNSPPR 720
DB 661 EHVGPASAVENIDKALATVTLQVGYTHFLTVKDOQGLSSTSTLTVAVKKENNSPPR 720
QY 721 ARAGRHVLPNNISITLDGSRSTDDQRIIVSYLMIRDGQSPAAGDVIDGSDHVALQLTN 780
DB 721 ARAGRHVLPNNISITLDGSRSTDDQRIIVSYLMIRDGQSPAAGDVIDGSDHVALQLTN 780
QY 781 LVEGYTHFLRVTDQSGASDSDTATVEVQDPKRGSLVELTLQVGVQLTQKOTLVRQ 840

DB 781 LVEGYTHFLRVTDQSGASDSDTATVEVQDPKRGSLVELTLQVGVQLTQKOTLVRQ 840
QY 841 LAVLLNVLDSIDIKVQKIRAHSDLSITVIVFYVQSRPPFKVLKAAEVARNLHMLSKERADF 900
DB 841 LAVLLNVLDSIDIKVQKIRAHSDLSITVIVFYVQSRPPFKVLKAAEVARNLHMLSKERADF 900
QY 901 LLPKVLAVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRIYIMDGESNCESIFVVT 960
DB 901 LLPKVLAVDTAGCLLKCSGHGCHDPLTKRCICSHLWMENLIQRIYIMDGESNCESIFVVT 960
QY 961 VLAFTLIVLTGGFTWLCICCCCKRQRTKIRKTKYTTILDNNDEQERMELRKYGIKHRST 1020
DB 961 VLAFTLIVLTGGFTWLCICCCCKRQRTKIRKTKYTTILDNNDEQERMELRKYGIKHRST 1020
QY 1021 EHNSSLMVSESEFSDSDQDTIFSRKMERGNPKVSMNGSIRNGASFYCSKDR 1072
DB 1021 EHNSSLMVSESEFSDSDQDTIFSRKMERGNPKVSMNGSIRNGASFYCSKDR 1072

RESULT 9

US-10-764-390-273
; Sequence 273, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 273
; LENGTH: 1053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-273

Query Match 98.3%; Score 5485; DB 16; Length 1053;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 20 CARQCSGRTYSNAVISPNLETTIRIMRVSHTPFVVDCTAACCDLSSCDLAWFEGRCYL 79
DB 1 CARQCSGRTYSNAVISPNLETTIRIMRVSHTPFVVDCTAACCDLSSCDLAWFEGRCYL 60
QY 80 VSCPHKENCPEKMGPIRSYLTFLVRPQRPQALLDYGMMNLRGSPSGIWGSDPEDIRK 139
DB 61 VSCPHKENCPEKMGPIRSYLTFLVRPQRPQALLDYGMMNLRGSPSGIWGSDPEDIRK 120
QY 140 DLXPLGKOWGLEMESEYDXYRELEKDLLQPSGKQPRGSAEYTDWGLLPGSEGAFNSV 199
DB 121 DLXPLGKOWGLEMESEYDXYRELEKDLLQPSGKQPRGSAEYTDWGLLPGSEGAFNSV 180
QY 200 GDSAPVAPARTQDPDELHYLNESASTPAPKLPERSVLLPLPTTSSGGEVLEKASQLOEQ 259
DB 181 GDSAPVAPARTQDPDELHYLNESASTPAPKLPERSVLLPLPTTSSGGEVLEKASQLOEQ 240
QY 260 SSNSGKEVLMPSHSLPPASLELSVTVVEKSPVLVTPGSTEHSTPTPTSAAPSESTPS 319
DB 241 SSNSGKEVLMPSHSLPPASLELSVTVVEKSPVLVTPGSTEHSTPTPTSAAPSESTPS 300
QY 320 ELPISTPTAPRTVKELTVSAGDNLITLPDNEVELKAFVAPPVETTYNENWLIHSHT 379


```

Db 301 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVADAPPVETTYNEMNLISHT 360
Qy 380 DYQGEIKQGHKOTLNLSQLSVGLVYFKVTVSSENAFGEFVNVTVPARRVNLPPVAVVS 439
Db 361 DYQGEIKQGHKOTLNLSQLSVGLVYFKVTVSSENAFGEFVNVTVPARRVNLPPVAVVS 420
Qy 440 POLQELTLPLTSALIDGOSTDDTEIVSYHWEENGPFIIEKTSVDSVPLRLSNLDPGNY 499
Db 421 POLQELTLPLTSALIDGOSTDDTEIVSYHWEENGPFIIEKTSVDSVPLRLSNLDPGNY 480
Qy 500 SFRLTVTDSGATNSTTAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGQSSDDHQI 559
Db 481 SFRLTVTDSGATNSTTAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGQSSDDHQI 540
Qy 560 VLYEWSLPGSGEGKHVVMQGVQTPYLHLISAMQEGDYTFQLKVTDSRQOSTAXVTIVQ 619
Db 541 VLYEWSLPGSGEGKHVVMQGVQTPYLHLISAMQEGDYTFQLKVTDSRQOSTAXVTIVQ 600
Qy 620 ENNRPPVAVAGDPKELIFPVESATLDGSSSDHGI VFYHWEHVRGSAVEMENIDKAIA 679
Db 601 ENNRPPVAVAGDPKELIFPVESATLDGSSSDHGI VFYHWEHVRGSAVEMENIDKAIA 660
Qy 680 TVTGLQVGYHFRLTVDKQOGLSSTLTAVVKKENNSPPRARAGRHVVLVLPNNSITLD 739
Db 661 TVTGLQVGYHFRLTVDKQOGLSSTLTAVVKKENNSPPRARAGRHVVLVLPNNSITLD 720
Qy 740 GSRSTDDQRIIVSYLWIRDOQSPAAGDVI DGSVHALQLTNLVEGYTFHLRVTDQSGAS 799
Db 721 GSRSTDDQRIIVSYLWIRDOQSPAAGDVI DGSVHALQLTNLVEGYTFHLRVTDQSGAS 780
Qy 800 DTDATATVEQDPKRSGLVELTLQVGVQLTEQRKDTLVRQLAVLNLVDSIKVQKIRA 859
Db 781 DTDATATVEQDPKRSGLVELTLQVGVQLTEQRKDTLVRQLAVLNLVDSIKVQKIRA 840
Qy 860 HSDLSSTVIVFYVQSRPPPKVLAEEVARNLHMLRSLKEKADFLFKVLVRVDTAGCLLKCSG 919
Db 841 HSDLSSTVIVFYVQSRPPPKVLAEEVARNLHMLRSLKEKADFLFKVLVRVDTAGCLLKCSG 900
Qy 920 HGHCPLTKRCICSHLWENLQRIYIWDGESNCEWSIFVYTVLAFTLVLTGQFTWLCIC 979
Db 901 HGHCPLTKRCICSHLWENLQRIYIWDGESNCEWSIFVYTVLAFTLVLTGQFTWLCIC 960
Qy 980 CCKRQKTKIRKTKYTIIDNDDEQRMELRPKYGIKHRSTEHNSLWSESEFSDQDT 1039
Db 961 CCKRQKTKIRKTKYTIIDNDDEQRMELRPKYGIKHRSTEHNSLWSESEFSDQDT 1020
Qy 1040 IFSREKMERGNPKVSMNGSIRNGASFSYCSKDR 1072
Db 1021 IFSREKMERGNPKVSMNGSIRNGASFSYCSKDR 1053

```

```

RESULT 10
US-10-764-390-274
; Sequence 274, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254P1d68 Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24

```

```

; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 1053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-274

Query Match      98.3%; Score 5485; DB 16; Length 1053;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 20 CARQCSGRTYSNAVISPNLETTIRIMRVSHFTFPVDCACCDLSSCDLAWPEGRCYL 79
Db 1 CARQCSGRTYSNAVISPNLETTIRIMRVSHFTFPVDCACCDLSSCDLAWPEGRCYL 60

Qy 80 VSCPHKENCERPKMGPIRSYLTFFVLRPVORPAQLDYGDMMLNRGSPGIGWGDSPEDIRK 139
Db 61 VSCPHKENCERPKMGPIRSYLTFFVLRPVORPAQLDYGDMMLNRGSPGIGWGDSPEDIRK 120

Qy 140 DLXFLGKDWGLEEMSEYKDDYRELEKDLLQPSGKQEPGSAEYTDWGLLPGEAGFNSV 199
Db 121 DLXFLGKDWGLEEMSEYKDDYRELEKDLLQPSGKQEPGSAEYTDWGLLPGEAGFNSV 180

Qy 200 GDSPAVPAETQDDPELHYLINESASTPAPKLPERSVLLPLPTTPSSGEVLEKEKASQLEQ 259
Db 181 GDSPAVPAETQDDPELHYLINESASTPAPKLPERSVLLPLPTTPSSGEVLEKEKASQLEQ 240

Qy 260 SSNSGSEKVLMPHSLPPASLELSVTVKESPVLTVPSTGSTEHSITPTPTSAAPSESTPS 319
Db 241 SSNSGSEKVLMPHSLPPASLELSVTVKESPVLTVPSTGSTEHSITPTPTSAAPSESTPS 300

Qy 320 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVADAPPVETTYNEMNLISHT 379
Db 301 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVADAPPVETTYNEMNLISHT 360

Qy 380 DYQGEIKQGHKOTLNLSQLSVGLVYFKVTVSSENAFGEFVNVTVPARRVNLPPVAVVS 439
Db 361 DYQGEIKQGHKOTLNLSQLSVGLVYFKVTVSSENAFGEFVNVTVPARRVNLPPVAVVS 420

Qy 440 POLQELTLPLTSALIDGOSTDDTEIVSYHWEENGPFIIEKTSVDSVPLRLSNLDPGNY 499
Db 421 POLQELTLPLTSALIDGOSTDDTEIVSYHWEENGPFIIEKTSVDSVPLRLSNLDPGNY 480

Qy 500 SFRLTVTDSGATNSTTAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGQSSDDHQI 559
Db 481 SFRLTVTDSGATNSTTAALIVNNAVDPYPVANAGPNHTITLPQNSITLNGQSSDDHQI 540

Qy 560 VLYEWSLPGSGEGKHVVMQGVQTPYLHLISAMQEGDYTFQLKVTDSRQOSTAXVTIVQ 619
Db 541 VLYEWSLPGSGEGKHVVMQGVQTPYLHLISAMQEGDYTFQLKVTDSRQOSTAXVTIVQ 600

Qy 620 ENNRPPVAVAGDPKELIFPVESATLDGSSSDHGI VFYHWEHVRGSAVEMENIDKAIA 679
Db 601 ENNRPPVAVAGDPKELIFPVESATLDGSSSDHGI VFYHWEHVRGSAVEMENIDKAIA 660

Qy 680 TVTGLQVGYHFRLTVDKQOGLSSTLTAVVKKENNSPPRARAGRHVVLVLPNNSITLD 739
Db 661 TVTGLQVGYHFRLTVDKQOGLSSTLTAVVKKENNSPPRARAGRHVVLVLPNNSITLD 720

Qy 740 GSRSTDDQRIIVSYLWIRDOQSPAAGDVI DGSVHALQLTNLVEGYTFHLRVTDQSGAS 799
Db 721 GSRSTDDQRIIVSYLWIRDOQSPAAGDVI DGSVHALQLTNLVEGYTFHLRVTDQSGAS 780

Qy 800 DTDATATVEQDPKRSGLVELTLQVGVQLTEQRKDTLVRQLAVLNLVDSIKVQKIRA 859
Db 781 DTDATATVEQDPKRSGLVELTLQVGVQLTEQRKDTLVRQLAVLNLVDSIKVQKIRA 840

Qy 860 HSDLSSTVIVFYVQSRPPPKVLAEEVARNLHMLRSLKEKADFLFKVLVRVDTAGCLLKCSG 919
Db 841 HSDLSSTVIVFYVQSRPPPKVLAEEVARNLHMLRSLKEKADFLFKVLVRVDTAGCLLKCSG 900

Qy 920 HGHCPLTKRCICSHLWENLQRIYIWDGESNCEWSIFVYTVLAFTLVLTGQFTWLCIC 979
Db 901 HGHCPLTKRCICSHLWENLQRIYIWDGESNCEWSIFVYTVLAFTLVLTGQFTWLCIC 960

Qy 980 CCKRQKTKIRKTKYTIIDNDDEQRMELRPKYGIKHRSTEHNSLWSESEFSDQDT 1039
Db 961 CCKRQKTKIRKTKYTIIDNDDEQRMELRPKYGIKHRSTEHNSLWSESEFSDQDT 1020

Qy 1040 IFSREKMERGNPKVSMNGSIRNGASFSYCSKDR 1072
Db 1021 IFSREKMERGNPKVSMNGSIRNGASFSYCSKDR 1053

```

```

Db 901 HGHCDPLTKRCICSHLWMENLIQRYIWDGSSNCWSIFYYTVTLAFTLIVITGGFTWLCIC 960
Qy 980 CCKRQKTKIRKTKYKTYTILNDMEQERMELEPKYGIKIRHSTERNSSLMVSESEFSDSDT 1039
Db 961 CCKRQKTKIRKTKYKTYTILNDMEQERMELEPKYGIKIRHSTERNSSLMVSESEFSDSDT 1020
Qy 1040 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1072
Db 1021 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1053

RESULT 11
US-10-764-390-7
; Sequence 7, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1063
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-7

Query Match 98.38; Score 5485; DB 16; Length 1063;
Best Local Similarity 99.74; Pred. No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 20 CARQKCEGRTYSNAVISPMLNLETRIMRVSHTFPVVDCTAACCDLSSCDLAWMPEGRCYL 79
Db 11 CARQKCEGRTYSNAVISPMLNLETRIMRVSHTFPVVDCTAACCDLSSCDLAWMPEGRCYL 70
Qy 80 VSCPHKENCBPCKMGPIRSYLTFTVLRVQRPQALLDYGDMLNRGSPSGIWGDSPEDIRK 139
Db 71 VSCPHKENCBPCKMGPIRSYLTFTVLRVQRPQALLDYGDMLNRGSPSGIWGDSPEDIRK 130
Qy 140 DLYFLGKDWGLEMSEYXDYRELEKDLQPSGKQEPGRGAETDYGWLLPGSEGAFFNSV 199
Db 131 DLYFLGKDWGLEMSEYXDYRELEKDLQPSGKQEPGRGAETDYGWLLPGSEGAFFNSV 190
Qy 200 GDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPPTTPSSGGVLEKEKASQLOEQ 259
Db 191 GDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPPTTPSSGGVLEKEKASQLOEQ 250
Qy 260 SSNSSGKEVLMPSHSLPPASLESVTVKESPVLTTPGSTEHSIPTPTSAAPSESTPS 319
Db 251 SSNSSGKEVLMPSHSLPPASLESVTVKESPVLTTPGSTEHSIPTPTSAAPSESTPS 310
Qy 320 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPVETTYNEMNLISHT 379
Db 311 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPVETTYNEMNLISHT 370
Qy 380 DYQGEIKQGHKOTLNLSQLSVGLVFKVTVSSNAFGEFVNVTVKPARVNLPPVAVVS 439
Db 371 DYQGEIKQGHKOTLNLSQLSVGLVFKVTVSSNAFGEFVNVTVKPARVNLPPVAVVS 430

```

```

Qy 440 POLQELTLPLTSALIDGSSQSDTTEIVSYHWEINGPFIEKTSVDSFVLRLSLNDPGNY 499
Db 431 POLQELTLPLTSALIDGSSQSDTTEIVSYHWEINGPFIEKTSVDSFVLRLSLNDPGNY 490
Qy 500 SFRLTVTSDGATNTSTAALIVNNAVDYPPVANAGPNHTITLPONSTTLNGNQSSDDHQI 559
Db 491 SFRLTVTSDGATNTSTAALIVNNAVDYPPVANAGPNHTITLPONSTTLNGNQSSDDHQI 550
Qy 560 VLYEWSLPGSEGKHVVMQVQTPYHLHLSAMQEGDYTFQLKVTDSSRQQSTAXTVTVIQQP 619
Db 551 VLYEWSLPGSEGKHVVMQVQTPYHLHLSAMQEGDYTFQLKVTDSSRQQSTAXTVTVIQQP 610
Qy 620 ENNRPPVAVAGDPKELIFPVESATLDGSSSDDDHGIYFYHWEHVRGSAVEMENIDKAIA 679
Db 611 ENNRPPVAVAGDPKELIFPVESATLDGSSSDDDHGIYFYHWEHVRGSAVEMENIDKAIA 670
Qy 680 TVTGLQVGYTHFRLTVKDOQGLSSTSLTVAVKKENNSPPRARAGGRHVLVLPNNSITLD 739
Db 671 TVTGLQVGYTHFRLTVKDOQGLSSTSLTVAVKKENNSPPRARAGGRHVLVLPNNSITLD 730
Qy 740 GSRSTDQRIVSYLWIRDOGSPAAGDVIDGSDHSHVALQTLNLVEGVYTFHLRVTDSSQAS 799
Db 731 GSRSTDQRIVSYLWIRDOGSPAAGDVIDGSDHSHVALQTLNLVEGVYTFHLRVTDSSQAS 790
Qy 800 DTDATVEVQDPDRKSGVELTLOQVGVQLTQEQKDTLVRQLAVLLNVLSDIKVQKIRA 859
Db 791 DTDATVEVQDPDRKSGVELTLOQVGVQLTQEQKDTLVRQLAVLLNVLSDIKVQKIRA 850
Qy 860 HSDLSTVIVFYVQSRPPFKVLAKEAVARNLHMLSKADFLFKVLRVDTAGCLLKCSG 919
Db 851 HSDLSTVIVFYVQSRPPFKVLAKEAVARNLHMLSKADFLFKVLRVDTAGCLLKCSG 910
Qy 920 HGHCDPLTKRCICSHLWMENLIQRYIWDGSSNCWSIFYYTVTLAFTLIVITGGFTWLCIC 979
Db 911 HGHCDPLTKRCICSHLWMENLIQRYIWDGSSNCWSIFYYTVTLAFTLIVITGGFTWLCIC 970
Qy 980 CCKRQKTKIRKTKYKTYTILNDMEQERMELEPKYGIKIRHSTERNSSLMVSESEFSDSDT 1039
Db 971 CCKRQKTKIRKTKYKTYTILNDMEQERMELEPKYGIKIRHSTERNSSLMVSESEFSDSDT 1030
Qy 1040 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1072
Db 1031 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKDR 1063

RESULT 12
US-10-764-390-10
; Sequence 10, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1063
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-10

```

```
Query Match      98.3%; Score 5485; DB 16; Length 1063;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 20 CARQCSGRTYNAVISPNLETTIRMRVSHFPVVDCTAACCDLSSCDLAWPFGRCYL 79
Db 11 CARQCSGRTYNAVISPNLETTIRMRVSHFPVVDCTAACCDLSSCDLAWPFGRCYL 70
Qy 80 VSCPKHCEPKMGPIRSYLTFLVRPVQRPQALLDYGDMMLNRGSPSGIWDSPEDIRK 139
Db 71 VSCPKHCEPKMGPIRSYLTFLVRPVQRPQALLDYGDMMLNRGSPSGIWDSPEDIRK 130
Qy 140 DLXFLGKDWGLEEMSEYKDDYRELEKOLLQPSGKQEPGRGSAEYTDWGLLPSEGFNFSV 199
Db 131 DLPFLGKDWGLEEMSEYKDDYRELEKOLLQPSGKQEPGRGSAEYTDWGLLPSEGFNFSV 190
Qy 200 GDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPTTPSSGEVLEKEKASQLQEQ 259
Db 191 GDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPTTPSSGEVLEKEKASQLQEQ 250
Qy 260 SSNSGKEVLMPSHSLPPASLELSVTVEKSPVLTVPSTGSTEHSIPTPTSAAPSESTPS 319
Db 251 SSNSGKEVLMPSHSLPPASLELSVTVEKSPVLTVPSTGSTEHSIPTPTSAAPSESTPS 310
Qy 320 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNWEWNLISHPT 379
Db 311 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNWEWNLISHPT 370
Qy 380 DYQGEIKQGHKQTLNLSQLSVGLYFVKVTVSSENAFGEFVNVTVPARRVNLPPVAVVS 439
Db 371 DYQGEIKQGHKQTLNLSQLSVGLYFVKVTVSSENAFGEFVNVTVPARRVNLPPVAVVS 430
Qy 440 POLQELTLP L T S A L I D G S Q S T D D T E I V S Y H W E E I N G P F I E E K T S V D S P V L R L S N L D P G N Y 499
Db 431 POLQELTLP L T S A L I D G S Q S T D D T E I V S Y H W E E I N G P F I E E K T S V D S P V L R L S N L D P G N Y 490
Qy 500 SFRLTVTDSGATNSTTAAALIVNNAVDYPPVANAGPNHTITLPQNSITLNGQSSDDHQI 559
Db 491 SFRLTVTDSGATNSTTAAALIVNNAVDYPPVANAGPNHTITLPQNSITLNGQSSDDHQI 550
Qy 560 VLYEWSLGFSGSGKHVMQGVQTPYLHLISAMQEGDYTFQLKVTDSRQSQSTAXVTVIQOP 619
Db 551 VLYEWSLGFSGSGKHVMQGVQTPYLHLISAMQEGDYTFQLKVTDSRQSQSTAXVTVIQOP 610
Qy 620 ENRPPVAVAGDPKELIIPVESATLDGSSSDDHGIVFYHWEHVHVRGSAVEMENIDKAIA 679
Db 611 ENRPPVAVAGDPKELIIPVESATLDGSSSDDHGIVFYHWEHVHVRGSAVEMENIDKAIA 670
Qy 680 TVTGLQVGYHPRLTVKDQOGLSSTLTVA VKENNSPPRARAGRHVLVLPNNSITLD 739
Db 671 TVTGLQVGYHPRLTVKDQOGLSSTLTVA VKENNSPPRARAGRHVLVLPNNSITLD 730
Qy 740 GSRSTDDQRIVSYLWIRDCQSPAAGDIVDGDSDHVALQTLNLVEGYTTFHLRVTDSQAS 799
Db 731 GSRSTDDQRIVSYLWIRDCQSPAAGDIVDGDSDHVALQTLNLVEGYTTFHLRVTDSQAS 790
Qy 800 DTDATVEVQDPKRSGLVELTQVGVGQLTEQRKDTLVRQLAVLNVLDSIDKVKQIRA 859
Db 791 DTDATVEVQDPKRSGLVELTQVGVGQLTEQRKDTLVRQLAVLNVLDSIDKVKQIRA 850
Qy 860 HSDLSVTIVFYVQSRPPPKVLAEEVARNLHRLSKEKADFLFKVLRVDTAGCLLKCSG 919
Db 851 HSDLSVTIVFYVQSRPPPKVLAEEVARNLHRLSKEKADFLFKVLRVDTAGCLLKCSG 910
Qy 920 HGHCDPLTKRCICSHLWMENLQRIYIWDGESNCWSIFVYTVLAFTLIVLTGFTWLCIC 979
Db 911 HGHCDPLTKRCICSHLWMENLQRIYIWDGESNCWSIFVYTVLAFTLIVLTGFTWLCIC 970
Qy 980 CCKRQKRTKIRKTKYITLDNMBQERMLRPKYIGIKHRSTEHNSLWVSESEFSDQDT 1039
Db 971 CCKRQKRTKIRKTKYITLDNMBQERMLRPKYIGIKHRSTEHNSLWVSESEFSDQDT 1030
```

```
Qy 1040 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKOR 1072
Db 1031 IFSREKMERGNPKVSMNGSIRNGASFYSYCSKOR 1063

RESULT 13
US-10-764-390-272
; Sequence 272, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge. Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PiD6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764.390
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 272
; LENGTH: 1063
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-272

Query Match      98.3%; Score 5485; DB 16; Length 1063;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 20 CARQCSGRTYNAVISPNLETTIRMRVSHFPVVDCTAACCDLSSCDLAWPFGRCYL 79
Db 11 CARQCSGRTYNAVISPNLETTIRMRVSHFPVVDCTAACCDLSSCDLAWPFGRCYL 70
Qy 80 VSCPKHCEPKMGPIRSYLTFLVRPVQRPQALLDYGDMMLNRGSPSGIWDSPEDIRK 139
Db 71 VSCPKHCEPKMGPIRSYLTFLVRPVQRPQALLDYGDMMLNRGSPSGIWDSPEDIRK 130
Qy 140 DLXFLGKDWGLEEMSEYKDDYRELEKOLLQPSGKQEPGRGSAEYTDWGLLPSEGFNFSV 199
Db 131 DLPFLGKDWGLEEMSEYKDDYRELEKOLLQPSGKQEPGRGSAEYTDWGLLPSEGFNFSV 190
Qy 200 GDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPTTPSSGEVLEKEKASQLQEQ 259
Db 191 GDSPAVPAETQDDPELHYLNESASTPAPKLPERSVLLPLPTTPSSGEVLEKEKASQLQEQ 250
Qy 260 SSNSGKEVLMPSHSLPPASLELSVTVEKSPVLTVPSTGSTEHSIPTPTSAAPSESTPS 319
Db 251 SSNSGKEVLMPSHSLPPASLELSVTVEKSPVLTVPSTGSTEHSIPTPTSAAPSESTPS 310
Qy 320 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNWEWNLISHPT 379
Db 311 ELPISTPTAPRTVKELTVSAGDNLIIITLPDNEVELKAFVAPAPPVETTYNWEWNLISHPT 370
Qy 380 DYQGEIKQGHKQTLNLSQLSVGLYFVKVTVSSENAFGEFVNVTVPARRVNLPPVAVVS 439
Db 371 DYQGEIKQGHKQTLNLSQLSVGLYFVKVTVSSENAFGEFVNVTVPARRVNLPPVAVVS 430
Qy 440 POLQELTLP L T S A L I D G S Q S T D D T E I V S Y H W E E I N G P F I E E K T S V D S P V L R L S N L D P G N Y 499
Db 431 POLQELTLP L T S A L I D G S Q S T D D T E I V S Y H W E E I N G P F I E E K T S V D S P V L R L S N L D P G N Y 490
Qy 500 SFRLTVTDSGATNSTTAAALIVNNAVDYPPVANAGPNHTITLPQNSITLNGQSSDDHQI 559
Db 491 SFRLTVTDSGATNSTTAAALIVNNAVDYPPVANAGPNHTITLPQNSITLNGQSSDDHQI 550
```

Qy 560 VLYEWSLPGSEGHVWVGQVOTPYLHLSAMQEGDYTFOLKVTDSRRQSTAXTVTVIOP 619
 Db 551 VLYEWSLPGSEGHVWVGQVOTPYLHLSAMQEGDYTFOLKVTDSRRQSTAXTVTVIOP 610
 Qy 620 ENNPPVAVAGPDELIPVESATLDGSSSSDDHGIYFVHWEHVRGSAVEMENIDKAI 679
 Db 611 ENNPPVAVAGPDELIPVESATLDGSSSSDDHGIYFVHWEHVRGSAVEMENIDKAI 670
 Qy 680 TVTGLQVTHFRITVTKDQGLSSTLTVAVKKENNSPPARAGRHVLVLPNNSITLD 739
 Db 671 TVTGLQVTHFRITVTKDQGLSSTLTVAVKKENNSPPARAGRHVLVLPNNSITLD 730
 Qy 740 GSRSTDORIVSYLWIRGQSPAGDVTDGSDHVALQTLNVEGVYTHFRVTDGAS 799
 Db 731 GSRSTDORIVSYLWIRGQSPAGDVTDGSDHVALQTLNVEGVYTHFRVTDGAS 790
 Qy 800 DTDATVEVQDPKSGELVELTQVGVQOLTEQRKDTLVRQAVLLNVLSDIKVKIRA 859
 Db 791 DTDATVEVQDPKSGELVELTQVGVQOLTEQRKDTLVRQAVLLNVLSDIKVKIRA 850
 Qy 860 HSDLTSTVTFVYQSRPPFKVLAABVARNLHMLRSKEKADFLFKVLKRVDTAGCLLCSG 919
 Db 851 HSDLTSTVTFVYQSRPPFKVLAABVARNLHMLRSKEKADFLFKVLKRVDTAGCLLCSG 910
 Qy 920 HGCDPLTKRCICSHLWENLIQRYIWDGSENCESIFVYTVLAFTLIVLTGGFTWLCIC 979
 Db 911 HGCDPLTKRCICSHLWENLIQRYIWDGSENCESIFVYTVLAFTLIVLTGGFTWLCIC 970
 Qy 980 CCKRQKTKRKTKTYTLDNMDOERMLPKYGIKHSRTEHNSLWVSEFSDQDT 1039
 Db 971 CCKRQKTKRKTKTYTLDNMDOERMLPKYGIKHSRTEHNSLWVSEFSDQDT 1030
 Qy 1040 IFSREKMERGNPKVSMNGSIRNGASFYSKSDR 1072
 Db 1031 IFSREKMERGNPKVSMNGSIRNGASFYSKSDR 1063

RESULT 14

US-11-097-143-16884
 ; Sequence 16884, Application US/11097143
 ; Publication No. US20050208558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Venter, J. Craig
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 ; TITLE OF INVENTION: DROSOPHILA GENES.
 ; FILE REFERENCE: CLO00728
 ; CURRENT APPLICATION NUMBER: US/11/097,143
 ; CURRENT FILING DATE: 2005-04-04
 ; PRIOR APPLICATION NUMBER: 60/157,832
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: 60/160,191
 ; PRIOR FILING DATE: 1999-10-19
 ; PRIOR APPLICATION NUMBER: 60/161,932
 ; PRIOR FILING DATE: 1999-10-28
 ; PRIOR APPLICATION NUMBER: 60/164,769
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: 60/173,383
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: 60/175,693
 ; PRIOR FILING DATE: 2000-01-12
 ; PRIOR APPLICATION NUMBER: 60/184,831
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/191,637
 ; PRIOR FILING DATE: 2000-03-23
 ; NUMBER OF SEQ ID NOS: 43008
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16884
 ; LENGTH: 1069
 ; TYPE: PRT
 ; ORGANISM: DROSOPHILA

US-11-097-143-16884

Query Match 26.0%; Score 1451.5; DB 20; Length 1069;
 Best Local Similarity 32.1%; Pred. No. 2,3e-80;
 Matches 373; Conservative 172; Mismatches 383; Indels 233; Gaps 35;
 Qy 8 LSSLLLVTTIAGCA-----RKCQSEGRYTSNAVISPNLETRIMRVSHST 51
 Db 8 ICNLLLLATAMSSAYADVTTONALLVSGKKHKTSPDNSVGGSS-ISPVLVCHKMLR--HV 64
 Qy 52 F-----PVVD-----CTAACCD-----LSSCDLAWWPEG 75
 Db 65 FENATPRDEQAGVFEEYKPPDAVEPLEEBAEYLNWCLQACCEKPRNGSSACNVLVFKA 124
 Qy 76 RCVLVSPCHKENCBPK--KMGPIRSYLTFLVRPV-----QRPALLDYGD 118
 Db 125 KCVHRCQSNACLPKLRVRM-PNEKVQMVLPVNLPLGDATWPLQLKAEAAKQNAEILPYDE 183
 Qy 119 MMLNRGSPGSGWGDSPEDIRKDLKFLGKDWGLEMSSEYXDD-----VRELEKDLLOPSGK 173
 Db 184 AALN-----FWKQP-----RRJSYLARN---QETPVTEDEDFPLADKRMQMIFQPDEN 229
 Qy 174 QEPRGSAEYTDWGLLPGESEGFNSSFVGDSPAVPAETQDDPELHYNESA-----STPA 226
 Db 230 -----DVLANEELGYDYDSNAKFTTCDMETPC 255
 Qy 227 PKUPERSVILLPLPTPSS-----GEVLEKEK-----ASQLOEBOSSNSSGKEV 268
 Db 256 PP-POQCV-----PLQPNNAVGRVCTCPGFFVWKNKORKCVMAAVPYSSYL---TSNEAGQOE 307
 Qy 269 LMPESHLPAPASLELSVTVVEKSPVLTVPSTGSTEHSIPTPTPSAAPSESTPSELPSPTTA 328
 Db 308 AAASENSPEVS-----TPPLKAEQN----- 327
 Qy 329 PRTVKELTVSAGDNLIITLPDNEVELKAFVAPAPPV-ETTYNVEWNLISHPT-DYQGBIK 386
 Db 328 ----KDIVSVMSK-EVRLPEQEVTLAAFTVPDQTSQTKYKYLWTLISQKPGMNGTIS 382
 Qy 387 QGHQTLNLSQLSVGLYVFKVTVSSEN-APGEGFNVTVKPARVNLPPVAVSPQLQEL 445
 Db 383 DQSKSKVKLSNLSEGLYTFKVTVDGNTGFEATANTVLPENRINQPPQVILSPREGII 442
 Qy 446 TLPLTSALIDGQSTDDTEIVSVHWEERINGPFIEKTSVDSPLVRLSNLDGNYSERLTV 505
 Db 443 RQPTTNAILDGSTDDDKITNHWHEVIGISGYPVLPEVNTLQDLTSPGNTFKLTV 502
 Qy 506 TDSGATNTTAAALIVNAVVDYPPVAPAGNHTITLPONSIITLNGNOSSDDHQVLYEWS 565
 Db 503 TDSNVNTTATIAVLKETDYAPVAGDAVILYLPNNVTLNGTASSDDHEIVAWEMT 562
 Qy 566 LGPGSEGHVWVGQVOTPYLHLSAMQEGDYTFOLKVTDSRRQSTAXTVTVIOPENRPP 625
 Db 563 KASDEAKAVDMQNTRTPYVQLSNLEEGMTFVLKTDGSGQSTAKVHVFPVKPTNSPP 622
 Qy 626 VAVAGPKELIFFVESATLDGSSSSDDHGIYFVHWEHVRGSAVEMENIDKAIATVGLQ 685
 Db 623 VAEAGSNTTSLPINWVLLNGSDKDDIGIKSYLWKLQSGNNVILKSNSSIANATSLT 682
 Qy 686 VGTVPRLTVKDOQGLSSTLTVAVKKENNSPPARAGRHVLVLPNNSITLDGSRSTD 745
 Db 683 LGLYEFELTVADENNTATDTTWKVIQVERNAAPAINAGGDHTVTLPATAYFNGSKSWD 742
 Qy 746 DQRIYSYLWIRGQSPAGDVTDGSDHVALQTLNVEGVYTHFRVTDGASDSDTDTAT 805
 Db 743 DLAVVKVLTWRDEHSLAAGVIVADTKPEVMILNVLVQGRVYFTLVSDDDQGLTSSDTSVS 802
 Qy 806 VEVPDPRKSGELVELTQVGVQOLTEQRKDTLVRQAVLLNVLSDIKVKIR--AHSDL 863
 Db 803 VNVRDEPKLNLVQMTPLPMGI SVLVQSELDSSVQKQLQLLGD-ENKIQVRELKYLDTHTD- 860
 Qy 864 STVIVFYVQSRPPFKVLAABVARNLHMLRSKEKADFLFKVLKRVDTAGCLLCSGHHGC 923
 Db 861 ATVLVIFYVNDGQG-KALDGLQVERQLRTQLQKQASILGAFV-DIRTSVCOSDCSGHSC 918

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 12, 2005, 10:33:12 ; Search time 3809 Seconds
(without alignments)
1960.176 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 5580
Sequence: 1 MAPPTGVLSLLLVITAGC.....VSMNGSIRNGASPSYCSKDR 1072

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Rgapop 6.0 , Rgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 8443130 seqs, 3482420727 residues

Total number of hits satisfying chosen parameters: 16886260

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/p/US10764390/runat_12102005_110204_22383/app_query.fasta_1.1223
-DB=published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10764390 @CGN 1 1 723 @runat_12102005_110204_22383
-NCPU=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5574	99.9	6791	20	US-10-764-390-2
2	5574	99.9	6791	20	US-10-764-390-4
3	5546.5	99.4	6797	20	US-10-764-390-270
4	5546.5	99.4	6797	20	US-10-764-390-271
5	5527	99.1	6991	20	US-10-764-390-6
6	5527	99.1	6991	20	US-10-764-390-269
7	1451.5	26.0	3345	24	US-11-097-143-16883
8	1314	23.5	2423	22	US-10-450-763-8985
9	1280.5	22.9	5775	24	US-11-097-143-16882
10	477	8.5	284	20	US-10-764-390-1
11	403	7.2	136990	19	US-10-717-597-68
12	397	7.1	1420	14	US-10-050-704-89
13	397	7.1	1420	19	US-10-798-512-89
14	241.5	4.3	3612	17	US-10-282-122A-13443
15	225	4.0	82027	21	US-10-278-698-1034
16	225	4.0	82027	21	US-10-278-698-1045
17	224	4.0	81940	9	US-09-759-508B-1
18	224	4.0	81940	10	US-09-360-706-1092
19	224	4.0	81940	10	US-09-873-319-724
20	224	4.0	81940	20	US-10-723-860-132
21	224	4.0	81940	20	US-10-656-873A-1
22	224	4.0	81940	22	US-10-756-149-144
23	221.5	4.0	110838	22	US-10-840-512-41
24	219.5	3.9	2379	21	US-10-875-518-3
25	214.5	3.8	3687	10	US-09-917-384-2
26	214.5	3.8	3687	10	US-09-917-383-2
27	214	3.8	23546	24	US-11-097-143-1223
28	213	3.8	16680	17	US-10-282-122A-38993
29	207	3.7	44014	24	US-11-097-143-1222
30	206.5	3.7	3586	17	US-10-282-122A-9979
31	204	3.7	8307	24	US-11-097-143-31982
32	204	3.7	11597	24	US-11-097-143-31981
33	202	3.6	2601	17	US-10-297-639-14
34	202	3.6	4951	22	US-10-450-763-23901
35	202	3.6	5403	21	US-10-956-157-1961
36	202	3.6	5403	22	US-10-450-763-13894
37	202	3.6	5403	22	US-10-450-763-29584
38	201.5	3.6	25165	14	US-10-114-170-39
39	201	3.6	3928	20	US-10-723-860-7627
40	199	3.6	2526	21	US-10-489-695-53
41	199	3.6	2537	21	US-10-489-695-52
42	199	3.6	17203	9	US-09-864-761-20867
43	199	3.6	17203	16	US-10-029-386-20814
44	199	3.6	182508	22	US-10-981-277-27
45	197	3.5	1555	19	US-10-470-390A-17

ALIGNMENTS

RESULT 1

US-10-764-390-2
; Sequence 2, Application US/10764390
; Publication No. US2004021421A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Sid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PD6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer

```

; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 6791
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (512)...(3730)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 286
; OTHER INFORMATION: s = g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 935
; OTHER INFORMATION: m = c or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 980
; OTHER INFORMATION: k = t or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2347
; OTHER INFORMATION: r = g or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3762
; OTHER INFORMATION: y = c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 3772; r = a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 3955; y = c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 4096; y = c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 4415; r = g or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 4519; r = g or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 4539; r = a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 4614; k = t or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 5184; s = g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 5528; k = t or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)

```

```

; OTHER INFORMATION: Pos: 5641; r = g or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 6221; y = t or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Pos: 6223; r = g or a
; US-10-764-390-2
Alignment Scores:
Pred. No.: 0 Length: 6791
Score: 5574.00 Matches: 1071
Percent Similarity: 99.91% Conservative: 0
Best Local Similarity: 99.91% Mismatches: 1
Query Match: 99.89% Indels: 0
DB: 20 Gaps: 0
US-10-764-390-3 (1-1072) x US-10-764-390-2 (1-6791)
Qy 1 MetalaProThrGlyValLeuSerSerLeuLeuValThrIleAlaGlyCys 20
Db 512 ATGGCGCCCCCAGGGTGTCTCTTCAATTGCTGCTGGTGACAAATTCAGGTGT 571
Qy 21 AlaArgLysGlnCysSerGlyArgThrTyrSerAsnAlaValIleSerProAsnLeu 40
Db 572 GCCGTAAAGCAGTCGACGGAGGAGGACATATCCATGTCAGTCATTTCCACTAAC 631
Qy 41 GluThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAla 60
Db 632 GAAACACCACGATCATGCGGGTGTCTCACACCTTCCCTGTCGTAGACTGCACGCG 691
Qy 61 CysCysAspLeuSerSerCysAspLeuAlaTrpPheGluGlyArgCysTyrLeuVal 80
Db 692 TGCTGTGACCTGTCCAGCTGTGACCTGGCTGGTTCGAGGGGCGCTGTCTACCTGG 751
Qy 81 SerCysProHisLysGluAsnCysGluProLysLysMetGlyProIleArgSerTyrLeu 100
Db 752 AGCTGCCCCCAAGAGAACTGTGAGCCCAAGAGATGGGCCCCCATCAGGTCTTATCTC 811
Qy 101 ThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMet 120
Db 812 ACTTTTGTGCTCCGGCTGTTTCAGAGGCTGCACAGCTGCTGGACATATGGGGACATG 871
Qy 121 LeuAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAspIleArgLysAsp 140
Db 872 CTGAACAGGGGCTCCCCCTCGGGGATCTGGGGGGAATCACCCTGAGGATATCAGAAAG 931
Qy 141 Leu***PheLeuGlyLysAspTrpGlyLeuGluMetSerGluTyr***AspAspTyr 160
Db 932 TTGMCCTTTCTAGGCNAAGATTGGGGCTAGAGGAGATGTCTGAGTACKAGATGACTAC 991
Qy 161 ArgGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAla 180
Db 992 CGGAGCTGGAGAAGGACCTCTTGCAACCCAGTGGCAAGCAGGAGGCCGCCAGAGGAGTGC 1051
Qy 181 GluTyrThrAspTrpGlyLeuLeuProGlySerGlyAlaPheAsnSerSerValGly 200
Db 1052 GAGTACACGGACTGGGGCTACTGCGGGCAGCGAGGGGGCTTCAACTCTCTGTGGA 1111
Qy 201 AspSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAsnGlu 220
Db 1112 GACAGTCTCGGTGCCAGCGAGAGCGAGGAGCCCTGAGTCCATTCCTGATGAG 1171
Qy 221 SerAlaSerThrProAlaProLysLeuProGluArgSerValLeuLeuProLeuProThr 240
Db 1172 TCGGCTTCAACCCCTGCCCAAACTCCCTCAGAGAGTGTGTCTCTCCCTTGGCGACT 1231
Qy 241 ThrProSerSerGlyGluValLeuGluLysGlyLysAlaSerGluGlnGlnGlnSer 260
Db 1232 ACTCCATCTTCAGGAGAGGTGTGTGAGAAAGAAAGGCTTCTCAGCTCCAGGACAAATCC 1291

```


Qy	261	SerAsnSerSerGlyLysGluValLeuMetProSerHisSerIleuProProAlaSerLeu	280
Db	1292	AGCAACAGCTCTGTGAAGAAGAGTTCTAATGCCTTCCCATAGTCTTCCTCGGGCAAGCCTG	1351
Qy	281	GluLeuSerSerValThrValGluLysSerProValLeuThrValThrProGlySerThr	300
Db	1352	GAGCTTAGCTTGATCACCGTGGAGAAAGCCAGTGTCTACAGTCAACCCTGGGGAGTACA	1411
Qy	301	GluHisSerIleProThrProThrSerAlaAlaProSerGluSerThrProSerGlu	320
Db	1412	GAGCACAGCATCCCAACACACTCCCACTAGCGCAGCCCCCTCTGAGTCCACCCCATCTGAG	1471
Qy	321	LeuProIleSerProThrThrAlaProArgThrVallysGluLeuThrValSerAlaGly	340
Db	1472	CTACCCATATCTCCTACACTGTCTCCAGGACAGTGAAGAACAATTACGGTATCGGGTGGGA	1531
Qy	341	AspAsnLeuIleIleThrLeuProAspAsnGluValGluLeuLysAlaPheValAlaapro	360
Db	1532	GATAACCTAATTATTACTTTACC CGCAATGAAGTTGA CTGAAGGCCCTTTGTTGGCCCA	1591
Qy	361	AlaProProValGluThrThrTyrrAsnTrpGluTrpAsnLeuIleSerHisProThrAsp	380
Db	1592	GCGCCACTGTAGAAA CAACCTCAACTATGAATGGAA TTTAATAAGCCACCCACACAGAC	1651
Qy	381	TyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerVal	400
Db	1652	TACCAAGGTGAAATATAAACCAAGGACACAAAGCAA ACTCTTAACTCTCTCTCAATTGTCGGTC	1711
Qy	401	GlyLeuTyrrValPheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPheVal	420
Db	1712	GGACTTTATGTCTTCAAAGTCAC TGTTTTCTAGTGA AAGCCCTTTGAGAAAGGATTTGTC	1771
Qy	421	AsnValThrVallysProAlaArqValVaAsnLeuProProValAlaValValSerPro	440
Db	1772	AATGTCACTGTAAAGCTGCCAGAAAGAGTCAACTGCCACTGTAGCAGTTGTTCTCCC	1831
Qy	441	GlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThr	460
Db	1832	CAACTGCAAGAGCTCACTTTGGCTTTTGACGTGAGCCCTCATTTGATGGCAGCAAAAGTACA	1891
Qy	461	AspAspThrGluIleValSerTyrrHisTrpClucluleAsnGlyProPheIleGluGlu	480
Db	1892	GATGATACTGAAATAGTGAGTTATCATTTGGAGAAATAAACGGGGCCCTTCATAGAAGAG	1951
Qy	481	LysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyVasnTyrrSer	500
Db	1952	AAGACTTCAGTTGACTCTCCCGCTTACGCTTGTCTAACTTGATCTCTGGTAACTATAGT	2011
Qy	501	PheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuile	520
Db	2012	TTCAAGTTGACTGTTACAGACTCGGACGGAGCCACTAACTCTACAACTGCAGGCCCTAATA	2071
Qy	521	ValAsnAsnAlaValAspTyrrProProValAlaAsnAlaGlyProAsnHisThrIleThr	540
Db	2072	GTGAACAAATGCTGTGGACATCCCAACCAAGTTGCTAATGCAGGACCAAAATCACACCAATAACT	2131
Qy	541	LeuProGlnAsnSerIleThrLeuAsnGlyVasnGlnSerSerAspAspHisGlnIleVal	560
Db	2132	TTGGCCCAAAACTCCATCACTTTGAATGGA AAAC CAGAGCAGTGA CATCACCA GATTGTCT	2191
Qy	561	LeuTyrrGluTrpSerLeuGlyProGlySerGluGlyLyshisValValMetGlnGlyVal	580
Db	2192	CTCTATGAGTGTCCCTGGGTCTGGGAGTGGAGGCCAAACATGTGGTCA TGACAGGAGTA	2251
Qy	581	GlnThrProTyrrLeuHisLeuSerAlaMetGlnGluGlyAspTyrrThrPheGlnLeuLys	600
Db	2252	CAGACGCCATACCTTTCATTTATCTGCAATGCAGAAAGAGAGATTATACATTTTCAGCTGAAG	2311
Qy	601	ValThrAspSerSerArqGlnGlnSerThrAla***ValThrValIleValGlnProGlu	620
Db	2312	GTGACAGATTCTTCAAGGCCAACGTCTACTCTGCTGTGTGACTGTGATTGTGTCCACGCCGAA	2371
Qy	621	AsnAsnArqProProValAlaValAlaGlyProAspLysGluLeuIlePheProValGlu	640

[illegible]

Db 3452 TGCAAAAGHCAAAAAAGGACTAAATACGAGAAAAAACAAGTACACCATCTCGGATAAC 3511
 QY 1001 MetAspGluGlnGluArgMetGluLeuArgProLysTyrGlyIleLysHisArgSerThr 1020
 Db 3512 ATGGATGAACAGAGAAAGATGAAGTGGAGCCCAATATGGTATCAAGCACCAGACACA 3571
 QY 1021 GluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIle 1040
 Db 3572 GAGCACAACTCCAGCCTGATGATGATATCCGAGTCTGAGTTTGACAGTGACAGGACACAATC 3631
 QY 1041 PheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerIleArg 1060
 Db 3632 TTCAGCCGAGAAAGATGAGAGAGGGAATCCAAAGGTTTCATGAATGGTTCCATCAGA 3691
 QY 1061 AsnGlyAlaSerPheSerTyrCysSerLysAspArg 1072
 Db 3692 AATGAGGCTTCCTTCAGTTATTCCTCAAGAGACAGA 3727

RESULT 2

US-10-764-390-4
 ; Sequence 4, Application US/10764390
 ; Publication No. US2004021412A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Jakobovits, Ava
 ; APPLICANT: Challita-Bid, Pia M.
 ; APPLICANT: Ge, Wangmao
 ; APPLICANT: Faris, Mary
 ; APPLICANT: Steven B. Kanner
 ; APPLICANT: Juan J. Perez-Villar
 ; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
 ; TITLE OF INVENTION: Entitled 254P1D6B Useful in Treatment and Detection of
 ; FILE REFERENCE: 51158-20081.00
 ; CURRENT APPLICATION NUMBER: US/10/764,390
 ; CURRENT FILING DATE: 2004-01-23
 ; PRIOR APPLICATION NUMBER: US60/442,526
 ; PRIOR FILING DATE: 2003-01-24
 ; NUMBER OF SEQ ID NOS: 277
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 6791
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (512)... (3730)
 US-10-764-390-4

Alignment Scores:
 Pred. No.: 0 Length: 6791
 Score: 5574.00 Matches: 1069
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 3
 Query Match: 99.89% Indels: 0
 DB: 20 Gaps: 0

US-10-764-390-3 (1-1072) x US-10-764-390-4 (1-6791)

QY 1 MetAlaProProThrGlyValLeuSerSerLeuLeuLeuValThrIleAlaGlyCys 20
 Db 512 ATGGCGCCCCACAGGTGTGCTCTCTTCATTGCTGCTGCTGGTGCACAAATGCAGGTGT 571
 QY 21 AlaArgLysGlnCysSerGluGlyArgThrTyrSerAsnAlaValIleSerProAsnLeu 40
 Db 572 GCCCGTAAGCAGTGCAGCGAGGGGAGACATATTCCAATGCAGTCATTTACCTAACTTG 631
 QY 41 GluThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAlaAla 60
 Db 632 GAAACCAACCAAGATCATCGGGGTCTCACACCTTCCCTGCTAGACTGCACGCGCGCT 691
 QY 61 CysCysAspLeuSerSerCysAspLeuAlaTrpPheGluGlyArgCysTyrLeuVal 80

Db 692 TGCTGTGACCTGTCCAGCTGTGACCTGTGCTGTGGTTCGAGGGCCGCTGCTACCTGGTGTG 751
 QY 81 SerCysProHisLysGluAsnGlyProLysLysMetGlyProIleArgSerTyrLeu 100
 Db 752 AGCTGCCCCCAACAAGAGAACTGTGAGCCCAAGAGATGGGCCCCCATCAGGTCTTATCTC 811
 QY 101 ThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMet 120
 Db 812 ACTTTTGTGCTCCGGCTGTTTCAGAGGCTGCACAGCTGCTGGACTATGGGGACATGATG 871
 QY 121 LeuAsnArgGlySerProSerGlyIleTyrPglyAspSerProGluAspIleArgLysAsp 140
 Db 872 CTGAACAGGGGCTCCCTCGGGATCTGGGGGAGACTCACCTGAGGATATCAGAAAGGAC 931
 QY 141 Leu***PheLeuGlyLysAspTyrGlyLeuGluMetSerGlyTyr***AspAspTyr 160
 Db 932 TTGGCTTTTCTAGGCAAGATTTGGGGCTCAGAGAGATGTCTGAGTACGCGAGATCAGTAC 991
 QY 161 ArgGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAla 180
 Db 992 CGGGAGCTGGAGAGGACCTCTTGCNACCAGTGGCAGCAGGAGCCACAGAGGAGTGCC 1051
 QY 181 GluTyrThrAspTyrGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGly 200
 Db 1052 GAGTACACGGACTGGGGCTACTGCGGGCAGCGAGGGGGCTTCAACTCTCTCTGTGGA 1111
 QY 201 AspSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAsnGlu 220
 Db 1112 GACAGTCTCGGGTGCCAGCGAGACGCGAGCAGACCCCTGAGCTCCATTACCTGAATGAG 1171
 QY 221 SerAlaSerThrProAlaProLysLeuProGluArgSerValLeuLeuProLeuProThr 240
 Db 1172 TCGGCTTCAACCCCTGCCCCAAAACCTCCCTCAGAGAGTGTGTGCTTCCCTTGGCGACT 1231
 QY 241 ThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGlnGlnSer 260
 Db 1232 ACTCCATCTTCAGGAGAGGTGTGAGAGAAAGAAAGGCTTCTCAGCTCCAGGAAACATCC 1291
 QY 261 SerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProProAlaSerLeu 280
 Db 1292 AGCAACAGCTCTGAAAGAGGTTCTAATGCTTCCCATAGTCTTCTCCGCGAAGCCTG 1351
 QY 281 GluLeuSerSerValThrValGluLysSerProValLeuThrValThrProGlySerThr 300
 Db 1352 GAGCTCAGCTCAGTCACTGAGAGAAAGCCAGTGTCTCAGTCCACCCCGGGAGTACA 1411
 QY 301 GluHisSerIleProThrProThrSerAlaAlaProSerGluSerThrProSerGlu 320
 Db 1412 GAGCACAGCATCCCAACACTCCCACTAGCGAGCGCCCTCTGAGTCCACCCCATCTGAG 1471
 QY 321 LeuProIleSerProThrThrAlaProArgThrValLysGluLeuLeuValSerAlaGly 340
 Db 1472 CTACCCCATATCTCTTACCACCTGCCAGGAGCAGTGAAGAACTTACCGGTATCGGCTGGA 1531
 QY 341 AspAsnLeuIleIleThrLeuProAspAsnGluValGluLeuLysAlaPheValAlaPro 360
 Db 1532 GATAACCTTAATTAACCTTTACCCGACAAATGAAGTTGAAGTGAAGCCCTTTGTTGGCCA 1591
 QY 361 AlaProProValGluThrThrTyrAsnTyrGluTrpAsnLeuIleSerHisProThrAsp 380
 Db 1592 GCGCCACCTGTAGAAAACAACCTACAACCTATGAATGGAATTTAATAAGCCACCCCCACAG 1651
 QY 381 TyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerVal 400
 Db 1652 TACCAGGTGAAATAAACCAAGGACACAGCAAACTCTTAACCTCTCTCAATTGTCGTC 1711
 QY 401 GlyLeuTyrValPheLysValThrValSerSerGluAsnAlaPheGlyGlyPheVal 420
 Db 1712 GGACTTTATGTTTCAAAAGTCACTGTTTCTAGTGAAGCCCTTTGGAGAGAGATTGTC 1771
 QY 421 AsnValThrValLysProAlaArgArgValAsnLeuProProValAlaValValSerPro 440

Db 1772 AATGTCACCTGTAAGCCTGCCAGAGAGTCAACCTGCCACCTGTAGCAGTTGTTCTCCC 1831
Qy 441 GlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuLeuLeuAspGlySerGlnSerThr 460
Db 1832 CAACCTGCAAGAGCTCACTTTGCTTTCAGCTCAGCCCTCATTTGATGGCAGCAAGTACA 1891
Qy 461 AspAspThrGluLeuValSerThrHisThrGluGluLeuLeuAsnGlyProPheLeuGluGlu 480
Db 1892 GATGATCTGAATAGTAGTGATTCATTTGGGAAGAAATAAAGCGGCCCTTCATAGAAGAG 1951
Qy 481 LysThrSerValAspSerProValLeuArgLeuSerAsnLeuLeuAspProGlyAsnTyrSer 500
Db 1952 AAGACTTCAGTTGACTCTCCGCTCTACCTTGTCTAACTTGTCTGGTAACTATAGT 2011
Qy 501 PheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuLeu 520
Db 2012 TTCAGTTGACTGTTACAGACTCGAGCGAGCCACTTAACCTACAACTGCAGCCCTAATA 2071
Qy 521 ValAsnAlaValAspTyrProProValAlaAlaAlaGlyProAsnHisThrIleThr 540
Db 2072 GTGAACCAATGCTGTGGACTACCCACCACTTGTCTAATGAGGAGCAAAATCAACCATTA 2131
Qy 541 LeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleVal 560
Db 2132 TTGCCCCCAAACTCCATCACTTTGAAATGGAAACCAAGAGCAGTGAAGATTGTC 2191
Qy 561 LeuTyrGluThrSerLeuGlyProGlySerGluGlyLysHisValValMetGlnGlyVal 580
Db 2192 CTCTATGAGTGGTCCCTGGGCTCTGGGAGTCAGGCGCAAAATGTTGGTTCATGAGGAGTA 2251
Qy 581 GlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAspTyrThrPheGlnLeuLys 600
Db 2252 CAGACGCCATACCTTCATTTATCTGCAATGCAAGAGAGATTAATACATTTCACTGAAG 2311
Qy 601 ValThrAspSerSerArgGlnGlnSerThrAla**ValThrValIleValGlnProGlu 620
Db 2312 GTGACAGATCTTCAAGGCAACAGCTACTCTGCTGTAGTACTGTGATGTGTCCAGCTGAA 2371
Qy 621 AsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuLeuPheProValGlu 640
Db 2372 AACAAATAGACCTCCAGTGGCTGTGGCGGCTCTGATMAAGAGCTGATCTTCCAGTGGAA 2431
Qy 641 SerAlaThrLeuAspGlySerSerSerSerAspAspHisGlyIleValPheTyrHisThr 660
Db 2432 AGTGTACCTCCCTGGATGGGAGCAGCAGCAGCATGACCAAGCAGTGTCTTCTACCACTGG 2491
Qy 661 GluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAlaThr 680
Db 2492 GAGCAGTCCAGAGGCCCCAGTGCAGTGGAGATGGAAATATTTGCAAGCAATAGCCACT 2551
Qy 681 ValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGly 700
Db 2552 GTGACTGCTCTCCAGTGGGACCTACCACTTCCGTTTGACAGTGAAGACCAAGCAGGGA 2611
Qy 701 LeuSerSerThrSerThrLeuThrValAlaValLysLysGluAsnAsnSerProProArg 720
Db 2612 CTGAGCAGCAGTCCACCTCACTGTGCTGTGAGAGAGGAAATAATAGTCTCTCCAGCA 2671
Qy 721 AlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAspGly 740
Db 2672 GCCCGGCTGTGGCAGACATGTTCTGTGCTTCCCAATAATTCATTACTTTGGATGGT 2731
Qy 741 SerArgSerThrAspAspGlnArgIleValSerTyrLeuThrPheArgAspGlyGlnSer 760
Db 2732 TCNAGTCTACTGATGACCAAGAAATTTGTCTCTATCTGTGATCCGGATGGCCAGAGT 2791
Qy 761 ProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsn 780
Db 2792 CCAGCAGCTGGAGATGTATCGATGGCTCTGACCAAGTGTGGCTCTGACGCTTACGAAT 2851
Qy 781 LeuValGluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAsp 800
Db 2852 CTGTGGAGGGGGTGTACATTTCCATTTCCAGTTGCGAGTCAACCGAGTCAGGGGGCTCGGAC 2911

Qy 801 ThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValGluLeu 820
Db 2912 ACAGACACTGCTCACTGTGGAGTGCAGCCAGACCTAGGAGAGTGGCTGTGGAGCTG 2971
Qy 821 ThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuValArgGln 840
Db 2972 ACCCTGCAGGGTGTGTGGCAGCTGACAGAGCAGCGGAAGGACACCTTGTGAGGAG 3031
Qy 841 LeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLysIleArgAlaHis 860
Db 3032 CTGGCTGTGCTGTGAACGTGCTGACTCGGACATTAAGGTCCAGAGAGATTCGGGCCAC 3091
Qy 861 SerAspLeuSerThrValIleValPheTyrValGlnSerArgProProPheLysValLeu 880
Db 3092 TCGATCTCAGCACCGGTGATGTTGTTTATGTATACAGAGCAGCGCCCTTTCAGAGTTCTC 3151
Qy 881 LysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAspPhe 900
Db 3152 AAAGCTGCTGAAGTGGGCCCGAAATCTGCACATGCGGCTCTCAAGAGGAGAGCTGACTTC 3211
Qy 901 LeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGlyHis 920
Db 3212 TTGCTTTTCAAGTCTTTGAGGGTTGATACAGCAGGTTGCTTCTGAAGTGTCTTGGCCAT 3271
Qy 921 GlyHisCysAspProLeuThrLysArgCysIleCysSerHisLeuThrMetGluAsnLeu 940
Db 3272 GGTCACTCGCAGCCCTCCACAAAGCGCTGCAATTTGCTCTCTACTTATGGATGGAGAACCTT 3331
Qy 941 IleGlnArgTyrIleThrAspGlyGluSerAsnCysGluThrSerIlePheTyrValThr 960
Db 3332 ATACAGGCTTATATCTGGATGGAGAGAGCAACTGTGAGTGGAGTATATTTCTATGTGACA 3391
Qy 961 ValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCysCys 980
Db 3392 GTGTTGGCTTTTACTCTTATTGCTAAACAGAGGTTTCACTTGGCTTTCATCTGCTGC 3451
Qy 981 CysLysArgGlnLysArgThrLysIleArgLysLysThrLysTyrThrIleLeuAspAsn 1000
Db 3452 TGCAAAAGACAAAAAAGAGCTAAATCAGGAAAAAACAAGTAGTACCATCTCGGATAAC 3511
Qy 1001 MetAspGluGlnGluArgMetGluLeuArgProLysTyrGlyIleLysHisArgSerThr 1020
Db 3512 ATGATGAAACAGGAAAGATGGAATCAGGCCCCAATATGTTATCAAGACCCGAGACACA 3571
Qy 1021 GluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIle 1040
Db 3572 GAGCACAACTCCAGCTGTATGTTATCCGAGTCTGAGTTTGACAGTGACGACGACACAAATC 3631
Qy 1041 PheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerIleArg 1060
Db 3632 TTCAGCCGAGAAAAAGATGGAGAGAGGGAATCCAAAGGTTTCCATGAATGGTTCCATCAGA 3691
Qy 1061 AsnGlyAlaSerPheSerTyrCysSerIysAspArg 1072
Db 3692 AATGAGCTTCTCTTCAGTTATTGCTCAAGAGACACA 3727

RESULT 3

US-10-764-390-270
; Sequence 270, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Fakis, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer

```

; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 270
; LENGTH: 6797
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-390-270

Alignment Scores:
Pred. No.: 0 Length: 6797
Score: 5546.50 Matches: 1069
Percent Similarity: 99.53% Conservative: 0
Best Local Similarity: 99.53% Mismatches: 3
Query Match: 99.40% Indels: 2
DB: 20 Gaps: 1

US-10-764-390-3 (1-1072) x US-10-764-390-270 (1-6797)

QY 1 MetAlaProThrGlyValLeuSerSerLeuLeuLeuValThrIleAla---G1 19
Db 514 ATGGCGCCCCCACAGGTGTCTCTTCAATGCTGCTGCTGGTGACAAATTCAGGTTGG 573

QY 19 yCysAlaArgLysGlnCysSerGluGlyArgThrYrSerAsnAlaValIleSerProAs 39
Db 574 TTGTGCCCGTAAGCAGTGCAGCGAGGGGAGGACATATTCCAATGCAGTCATTTTCACTAA 633

QY 39 nLeuGluThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAl 59
Db 634 CTTGGAAACCAACAGNATCATGGGGTGCTCACACCTTCCCTGCTGTAGACTGCACGGC 693

QY 59 alaCysCysAspLeuSerSerCysAspLeuAlaTrpTrpPheGluGlyArgCysTyrlie 79
Db 694 CGCTTGTCTGACCTGTCCAGCTGTGACCTGGCCTGGTGGTTCGAGGGCGCTGTACTCT 753

QY 79 uValSerCysProHisLysGluAsnCysGluProLysLysMetGlyProIleArgSerTy 99
Db 754 GGTGAGCTGCCCCCAAAAGAGAACTGTGAGGCCAAGAGATGGGGCCCCCATCAGTCTTTA 813

QY 99 rLeuThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrglyAspMe 119
Db 814 TCTCATTTTGTCTCCGCCCTGTTTCAGAGGCTGCACAGCTGCTGGACTATGGGGACAT 873

QY 119 tMetLeuAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAspIleArgly 139
Db 874 GATGCTGAACAGGGGCTCCCCCTCGGGGATCTGGGGGACTCACCTGAGGATATCAGAAA 933

QY 139 sAspLeu***PheLeuGlyIleAspTrpGlyLeuGluGluMetSerGluTyri**AspAs 159
Db 934 GGACTTGGCCCTTTCTAGGCAAAAGATTGGGGCTTAGAGGAGATGTCTGAGTACTCAGATGA 993

QY 159 pTyArgGluLeuLeuLysAspLeuGlnProSerGlyLysGlnGluProArgGlyse 179
Db 994 CTACCGGGAGCTGGAGAGGACCTCTTTGCAACCCAGTGGCAAGCAGGAGGCCACAGGGGAG 1053

QY 179 rAlaGluTyThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerVa 199
Db 1054 TGCCGAGTACAGGACTGGGGCTACTGCGGGCAGCGAGGGGGCCCTTCAACTCTCTGT 1113

QY 199 lGlyAspSerProAlaValProAlaGluThrGlnInAspProGluLeuHisTyrlieAs 219
Db 1114 TGAGACAGTCTTCGCGGTGCCACGGAGACGACAGGAGCCCTGAGCTCCATTTACCTGAA 1173

QY 219 nGluSerAlaSerThrProAlaProLysLeuProGluArgSerValLeuLeuProLeuPr 239
Db 1174 TGAGTGGGTTTCAACCCCTTGGCCCAAACTCCCTGAGAGAGTGTGTCTTCCCTTGGC 1233

QY 239 oThrThrProSerSerGlyValLeuGluLysGluLysAlaSerGlnLeuGlnGluGl 259

```

```

Db 1234 GACTACTCCATCTTTCCAGGAGAGGTGTTGGAGAAAAGAAAGGCTTCTCAGCTCCAGGAACA 1293
QY nSerSerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProProAlase 279
Db 1294 ATCCAGCAACAGCTCTGGAAAAGAGGTTCTTAATGCTTCCCATAGTCTTCTTCCCGGCAAG 1353
QY rLeuGluLeuSerSerValThrValGluLysSerProValLeuThrValThrProGlyse 299
Db 1354 CCTGGAGCTCAGCTCAGTCACCGTGGAGAAAAGCCAGTGTCTCAGTCACTCCCGGGGAG 1413
QY rThrGluHisSerIleProThrProThrProThrSerAlaAlaProSerGluSerThrProSe 319
Db 1414 TACAGACACAGCATCCCAACACCTCCACCTAGCGCAGCCCTCTGTAGTCCACCCCATC 1473
QY rGluLeuProIleSerProThrThrAlaProArgThrValLysGluLeuThrValSerAl 339
Db 1474 TGAGCTACCCATATCTCTTACCACTGTCTCCAGACAGTGAAGNACTTACGGTATCGGC 1533
QY aGlyAspAsnLeuIleIleThrLeuProAspAsnGluValGluLeuLysAlaPheValAl 359
Db 1534 TGGAGATAACCTAATTTATACTTTTACCGCAATGAAGTTGAACCTGAAGGCTTTGTTC 1593
QY aProAlaProProValGluThrThrThrYrAsnTyrgluTrpAsnLeuIleSerHisProTh 379
Db 1594 GCCAGCGCCACTGTAGAAACACCTACACTATGAATGGAATTTAATAAGCCACCCAC 1653
QY rAspTyrglnGlnGlyLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSe 399
Db 1654 AGACTTACCAGGTGAATATAAACCAAGGACACAGCAAACTCTTAACCTCTCTCAATTGTC 1713
QY rValGlyLeuTyThrValPheLysValThrValSerSerGluAsnAlaPheGlyGluGlyph 419
Db 1714 CGTCGGACTTTATGTCTTCAAGTCACCTGTTCTAGTGAAACGCTTTGGAGAGGATT 1773
QY eValAsnValThrValLysProAlaArgArgValAsnLeuProProValAlaValValSe 439
Db 1774 TGTCAATGTCTACTGTAAAGCTGCGCAGAGAGTCAACCTGCGCACCTGTAGCAGTTGTTC 1833
QY rProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuLeuAspGlySerGlnSe 459
Db 1834 TCCCACTGCAAGAGCTCATTGTCCTTTCAGCTGAGCTCAGCTCATTTGATGGCAGCAGCAAG 1893
QY rThrAspAspThrGluIleValSerTyrglyTrpGluLeuLeuAsnGlyProPheIleGl 479
Db 1894 TACAGATGATCTGAAATAGTGTATCATTTGGAGAGAAATAAACGGCGCTTTCATAGA 1953
QY uGluLysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTy 499
Db 1954 AGAGAAGACTTTCAGTTGACTCTCCGCTTGTACGCTTGTCTAACCTTGATCTCGTGAACCTA 2013
QY rSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAla 519
Db 2014 TAGTTTCAGGTTGACTGTTACAGACTCGAGCGGACCACTAACTCTACACTGCGAGCCCT 2073
QY uIleValAsnAsnAlaValAspTyThrProValAlaAlaAsnAlaGlyProAsnHisThrI 539
Db 2074 AATAGTGAACAATGCTGTGGACTACCCACCAAGTTGCTAATGCAGGACCAAAATCACACAT 2133
QY eThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnI 559
Db 2134 AACTTTGCCCCAAAACCTCATCTTTTGAATGGAAACCCAGAGCAGTGCATCACAGAT 2193
QY eValLeuTyThrGluTrpSerLeuGlyProGlySerGluGlyLysHisValValMetGlnGl 579
Db 2194 TGTCTCTATGAGTGGTCCCTGGGTCTTGGGAGTGAGGGGCAACATGTGTTCTATCGAGGG 2253
QY yValGlnThrProTyThrLeuHisLeuSerAlaMetGlnGluGlyAspTyThrPheGlnLe 599
Db 2254 AGTACAGACGCATACCTTCACTTTATCTGCAATGCAGGAGGAGATTATACATTTTCAGCT 2313
QY uLysValThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnPr 619
Db 2314 GAAGGTGACAGATTCTTCAAGGCAACAGTCTACTGCTGTGGTGAAGTGTGATTGTCCAGGCC 2373

```

```
QY 619 oGluAsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProVa 639
Db 2374 TGAACAACATAGACCTCCAGTGGCTGTGGCGGCCCTGATAAAGAGCTGATCTTCCCAAGT 2433
QY 639 lGluSerAlaThrLeuAspGlySerSerSerSerSerSerSerSerSerSerSerSerSer 659
Db 2434 GGAAAGTCTACCTCGATGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2493
QY 659 sTriPGLuHisValArgGlyProSerAlaValAlaGluMetGluAsnIleAspLysAlaIleAl 679
Db 2494 CTGGGAGCACGTCAGAGGCCCCAGTGCAGTGGAGATGGAATAATTTGACAAAGCAATAGC 2553
QY 679 aThrValThrGlyLeuGlnValGlyThrTyxHisPheArgLeuThrValLysAspGlnGl 699
Db 2554 CACTGTGACTGGTCTCCAGGTGGGACCTACCCACTTCCGTTTGGACAGTGAAGACCAAGCA 2613
QY 699 nGlyLeuSerSerThrSerThrLeuThrValAlaValLysLysGluAsnAsnSerProPr 719
Db 2614 GGGACTGAGCAGCAGCTCCACCTCAGTGTGGCTGTGAAGAGGAAATAATATGTCCTCC 2673
QY 719 oArgAlaArgAlaGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAs 739
Db 2674 CAGAGCCCGGCTGGTGGCAGACATGTTCTTGTGCTTCCCAATAATTCATTACTTTGGA 2733
QY 739 pGlySerArgSerThrAspAspGlnArgIleValSerTyxIleuTrpIleArgAspGlyGl 759
Db 2734 TGGTTCAGAGTCTACTGTATGACCAAGAAATTTGTCTCTATCTGTGGATCCGGGATGGCCA 2793
QY 759 nSerProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuTh 779
Db 2794 GAGTCCAGCAGCTGGAGATGTCATCGATGGCTGTGACCAACAGATGGCTCTGACAGCTTAC 2853
QY 779 rAsnLeuValGluGlyValTyxThrPheHisLeuArgValThrAspSerGlnGlyAlaSe 799
Db 2854 GAACTGGTGGAGGGGTGTACACTTTCCACTTGGAGTCCACGACAGTCCAGGGGGCTC 2913
QY 799 rAspThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValGl 819
Db 2914 GGACACAGACACTGCACTGTGGAGTGCAGCCAGACCTAGGAAGAGTGGCTGGTGGGA 2973
QY 819 uLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuValAr 839
Db 2974 GCTGACCTCGAGGTGTGTGGTGGCAGCTGACAGAGCGGAGGACACCTTTGTGAG 3033
QY 839 gGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLysIleArgAl 859
Db 3034 GCAGCTGGCTGTGCTGACGTGCTGAGCTCGGACATTAAGGTCCAGAGAGATTCGGGC 3093
QY 859 aHisSerAspLeuSerThrValIleValPheTyxValGlnSerArgProProPheLysVa 879
Db 3094 CCACTCGGATCTCAGCACCGTGATTTGTTTATGTATGATACAGAGCGCGCCCTTTCAAGGT 3153
QY 879 lLeuLysAlaAlaGluValAlaAArgAsnLeuHisMetArgLeuSerLysGluLysAlaAs 899
Db 3154 TCTCAAGCTGCTGAAGTGGCCCGGAAATCTGCATGCGGCTCTCAAAAGGAGAGGCTGA 3213
QY 899 pPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGl 919
Db 3214 CTTCCTGCTTTTCAAGGTCTTGAGGGTTGATACAGCAGGTGGCTTCTGAAGTGTCTGG 3273
QY 919 yHisGlyHisCysAspProLeuThrLysArgCysIleCysSerHisLeuTrpMetGluAs 939
Db 3274 CCATGGTCACTGCGACCCCTCACAAAGCGCTGCAATTTGCTCTCACTTATGGATGGAGAA 3333
QY 939 nLeuIleGlnArgTyxIleTrpAspGlyGluSerAsnCysGluTrpSerIlePheTyxVa 959
Db 3334 CCTATACAGCGTTATATCTGGGATGGAGAGCAACTGTGAGTGAGATATATCTATGT 3393
QY 959 lThrValLeuAlaPheThrLeuIleValLeuThrGlyLysPheThrTrpLeuCysIleCy 979
Db 3394 GACAGTGTGGCTTTACTCTTATTTGCTTAACAGGAGTTTCACTTGGCTTTGCACTG 3453
```

```
QY 979 sCysCysLysArgGlnLysArgThrLysIleArgLysLysThrLysTyxThrIleLeuAs 999
Db 3454 CTGCTGCAAAAGACAAAAAAAGGACTAAATCAGGAAAAAAACAAAGTACACCATCTTGGGA 3513
QY 999 pAsnMetAspGluGlnGluArgMetGluLeuArgProLysTyxGlyIleLysHisArgSe 1019
Db 3514 TAACATGATGAACAGAGGAAGATGGAACCTGAGGCCCAATATGTTATCAAGCACCAGAG 3573
QY 1019 rThrGluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspTh 1039
Db 3574 CACAGACACAACCTCCAGCCTGATGGTATCCGAGTCTGAGTTTGACAGTGACCCAGGACAC 3633
QY 1039 rIlePheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerIl 1059
Db 3634 AATCTTCAGCCGAGAAAGATGGAGAGAGGAAATCCAAAGGTTTTCCATGAATGTTCCAT 3693
QY 1059 eArgAsnGlyAlaSerPheSerTyxSerSerLysAspArg 1072
Db 3694 CAGAAATGGAGCTTCTTTCAGTTATTGCTCAAGGACAGA 3733

RESULT 4
US-10-764-390-271
; Sequence 271, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764, 390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271
; LENGTH: 6797
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-390-271

Alignment Scores:
Pred. NO.: 0 Length: 6797
Score: 5546.50 Matches: 1069
Percent Similarity: 99.53% Conservative: 0
Best Local Similarity: 99.53% Mismatches: 3
Query Match: 99.40% Indels: 2
DB: 20 Gaps: 1

US-10-764-390-3 (1-1072) x US-10-764-390-271 (1-6797)
QY 1 MetAlaProThrGlyValLeuSerSerLeuLeuValThrIleAla---Gl 19
Db 514 ATGGCGCCCCCAGAGGTGCTCTCTTCATTGCTGTGTGGTGGACAAATTCAGTTTGG 573
QY 19 YCysAlaArgLysGlnCysSerGluGlyArgThrTyxSerAsnAlaValIleSerProAs 39
Db 574 TTGTGCCCGTAAGCAGTGCAGCGAGGGAGGACATATCCAAATGAGTCATTTACCTAA 633
QY 39 nLeuGluThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAl 59
Db 634 CTTGGAAACCAACCAAGAAATCATCGGGTGTCTCACACCTTCCCTGCTGAGACTGCACGCG 693
QY 59 aAlaCysCysAspLeuSerSerCysAspLeuAlaTyrPheGluGlyArgCysTyxLe 79
```

Db 694 CGCTTGCTGTGACCTGTCCAGCTGTGACCTGGCTGGTGGTTTCGAGGGCCGCTGCTACCT 753
Qy 79 uValSerCysProHisLysGluAsnCysGluProLysMetGlyProIleArgSerTy 99
Db 754 GGTGAGCTGCCCCCAAGAGAACTGTGAGCCCAAGAAAGATGGGCCCATCAGGTCTTA 813
Qy 99 rLeuThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMe 119
Db 814 TCTCACTTTTGTGCTCCGSCCTGTTTCAGAGGCCTGCACAGCTGCTGGACTATGGGGACAT 873
Qy 119 tMetLeuAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluLeuAspIleArgIly 139
Db 874 GATGCTGAACAGGGGCTCCCCCTCGGGATCTGGGGGACTCACCTGAGGATATCAGAAA 933
Qy 139 eAspLeu***PheLeuGlyLysAspTrpGlyLeuGluMetSerGluTyr***AspAs 159
Db 934 GGACTTGCCCTTCTTAGGGCAAGATTTGGGGCTTAGAGGAGATGTCTGAGTACTCAGATGA 993
Qy 159 pTyrArgGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlyse 179
Db 994 CTACCGGGAGCTGGAGAGGACCTCTTGCACCCAGTGGCAAGCAGGAGGCCAGAGGGAG 1053
Qy 179 rAlaGluTyrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerVa 199
Db 1054 TGCCGAGTACACGGACTGGGGCTTACTGCGGGCAGCGAGGGGGCTTCAACTCCTCTGT 1113
Qy 199 lGlyAspSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAs 219
Db 1114 TGGAGACAGTCTCGGTGCCAGCGGAGCGGACGAGGACCCTGAGCTCCATTACTGAA 1173
Qy 219 nGluSerAlaSerThrProAlaProLysLeuProGluArgSerValLeuLeuProLeuPr 239
Db 1174 TGAGTCCGCTTCAACCCCTGCCCAAACTCCTCGAGAACTGTGTGCTTCCCTTGC 1233
Qy 239 oThrThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGlnGlu 259
Db 1234 GACTACTCCATCTTCAGGAGAGGTGTGGAGAAAGAAAGGCTTCTCAGCTCCAGGAACA 1293
Qy 259 nSerSerAsnSerGlyLysGluValLeuMetProSerHisSerLeuProProAlaSe 279
Db 1294 ATCCAGCAACAGCTCTGGAAGAGAGGTCTTAATGCTTCCCAATAGTCTTCCCTCCGGCAAG 1353
Qy 279 rLeuGluLeuSerSerValThrValGluLysSerProValLeuThrValThrProGlyse 299
Db 1354 CTGGAGCTCAGCTCAGTCACCGTGAGAAAGCCAGTGCTCACAGTCACCCCGGGAG 1413
Qy 299 rThrGluHisSerIleProThrProThrSerAlaProSerGluSerThrProSe 319
Db 1414 TACAGAGCACAGCATCCCCAACACCTCCCACTAGCGCAGCCCCCTCTGAGTCCACCCCATC 1473
Qy 319 rGluLeuProIleSerProThrThrAlaProArgThrValLysGluLeuThrValSerAl 339
Db 1474 TGAGCTACCCTATCTCTTACCCTGCTCCAGGACAGTGAAGAACTTTACCGTATCGGC 1533
Qy 339 aGlyAspAsnLeuIleIleThrLeuProAspAsnGluValGluLeuLysAlaPheValAl 359
Db 1534 TGGAGATAACCTAAATTAATCTTACCCGACAATGAAGTTGAAGTGAAGGCTTTGTTGC 1593
Qy 359 aProAlaProProValGluThrThrTyrAsnTyrGluTrpAsnLeuIleSerHisProTh 379
Db 1594 GCCAGCGCCACCTGTAGAACAACCTTACAACATATGAATGGAATTTAATAAGCCACCCAC 1653
Qy 379 rAspTyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSe 399
Db 1654 AGACTACCAAGTGAATTAATAACAAGACACAAAGCAAACTTTAACTCTCTCAATTGTC 1713
Qy 399 rValGlyLeuTyrValPheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPh 419
Db 1714 CGTCGGACTTTATGCTCTCAAGTCACTGTTTCTAGTGAAGAACGCCCTTTGGAGAGGATT 1773
Qy 419 eValAsnValThrValLysProAlaArgValAsnLeuProProValAlaValAlSe 439
Db 1774 TGTCAATGTCACTGTTAAGCCTGCCAGAAAGAGTCAACCTGCCACCTGTAGCAGTGTGTTTC 1833

Qy 439 rProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSe 459
Db 1834 TCCCAACTGCAAGAGCTCACTTTGCCCTTGACCTCAGCCCTCATTTGATGCGCAAG 1893
Qy 459 rThrAspAspThrGluIleValSerTyrHisTrpGluGluLeuAsnGlyProPheIleGl 479
Db 1894 TACAGATGATACTGAAATAGTGTATCATTTGGGAAGAAATAAACGGGCCCTTCATAGA 1953
Qy 479 uGluLysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTy 499
Db 1954 AGAAGAGACTTCAGTTGACTCTCCCGCTTACGCTTGCTAACTTGATCTCTGGTAATA 2013
Qy 499 rSerPheArgLeuThrValThrAspSerAspGlyValaThrAsnSerThrThrAlaAlaLe 519
Db 2014 TAGTTTCAGGTGACTGTTACAGACTCGAGCGGAGCCACTAACTCTCAACTGCAGCCCT 2073
Qy 519 uIleValAsnAsnAlaValAspTyrProProValAlaAsnAlaGlyProAsnHisThrIl 539
Db 2074 AATAGTGAACAATGCTGTGACTACCCACCACTGCTAATGCAGGACCAAAATCACACCAT 2133
Qy 539 eThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIl 559
Db 2134 AACTTTGCCCCCAAACTCCATCACTTTGAATGGAAACCAGAGCAGTGACATCACAGAT 2193
Qy 559 eValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHisValValMetGlnGl 579
Db 2194 TGTCTCTCTAGTGGTCCCTGGGTCTCTGGGAGTGAGGGCAACATGTTGGTCATCAGGG 2253
Qy 579 yValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAspTyrThrPheGlnLe 599
Db 2254 AGTACAGCGCATACCTTCACTTTATCTGCAATGCAGAGGAGATTTATACATTTTCAGCT 2313
Qy 599 uLysValThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnPr 619
Db 2314 GAAGGTGACATTTCTCAGGCAACAGTCTACTGCTGTGGTGACTGTGATTTGTCCAGCC 2373
Qy 619 oGluAsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProVa 639
Db 2374 TGAACAACATAGACTCCAGTGGCTGTGGCCGGCCCTGATTAAGAGCTGATCTTCCAGT 2433
Qy 639 lGluSerAlaThrLeuAspGlySerSerSerSerAspAspHisGlyIleValPheTyrHi 659
Db 2434 GGAAGTGTCTACCTCGATGGAGCAGCAGCAGCGCATCACCCAGGCATTTGTTCTTACCA 2493
Qy 659 sTrpGluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAl 679
Db 2494 CTGGAGCACGTGAGAGGCCCCAGTGCAGTGGAGATGGAAATATTGAAGAAGCAATAGC 2553
Qy 679 aThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrValLysAspGlnGl 699
Db 2554 CACTGTGACTGTTCTCCAGGTGGGACCTACCACCTTCCGTTTGACAGTGAAGACCA 2613
Qy 699 nGlyLeuSerSerThrSerThrLeuThrValAlaValLysLysGluAsnAsnSerProPr 719
Db 2614 GGGACTGAGCAGCAGCTCCACCTCACTGTGGCTGTGAAGAGGAAAAATAATAGTCTCTCC 2673
Qy 719 oArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAs 739
Db 2674 CAGAGCCCGGGCTGTGGCAGACATGTTCTGTGCTTCCCAATAATTCCATTACTTTGA 2733
Qy 739 pGlySerArgSerThrAspAspGlnArgIleValSerTyrLeuTrpIleArgAspGlyGl 759
Db 2734 TGGTTCAAGTCTACTGATGACCAAGAAATTTGTCTTCTATCTGTGATCCGGATGGCCA 2793
Qy 759 nSerProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuTh 779
Db 2794 GAGTCCAGCAGCTGGAGATGTCTCGATGGCTCTGACCACAGTGTGGCTCTCAGCTTAC 2853
Qy 779 rAsnLeuValGluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSe 799
Db 2854 GAATCTGGTGGAGGGGGTGCACATTTTCCATTTGCGAGTCCCGACAGTCAGGGGGCCTC 2913

Qy	799	rAspThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValGI	819
Db	2914	GGACACAGACACTGCCACTGTGGAAGTGCAGCCAGACCCCTAGGAAGAGTGGCTCGTGGGA	2973
Qy	819	uLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuValAr	839
Db	2974	GCTGACCCTCGAGGTTGGTGTGGCAGCTGCACAGACGACGCGAGGACACCTTTGTGAG	3033
Qy	839	gGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLysIleArgAl	859
Db	3034	GCAGCTGGCTGTGCTGTAACGCTGTGGACTCGGACATTAAAGTCTCAGAAGATTCCGGGC	3093
Qy	859	aHisSerAspLeuSerThrValIleValPheTyrValGlnSerArgProProPheLysVa	879
Db	3094	CCACTCGGACTCAGACCCGTGATTGTGTTTTATGTACAGACGAGCGCGCTTCCTCAAGGT	3153
Qy	879	lLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAs	899
Db	3154	TCTCAAGCTGCTGAAGTGGCCCCGAAATCTGCACATGCGGCTCTCAAAGGAGAAGGCTGA	3213
Qy	899	pPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGI	919
Db	3214	CTTCTTGCTTTTCAAGCTCTTGAGGGTTGATACAGCAGGTTGCTTCTCGAAGTGTCTTGG	3273
Qy	919	yHisGlyHisCysAspProLeuThrLysArgCysIleCysSerHisLeuTrpMetGluAs	939
Db	3274	CCATGTCACATGGACCCCTCCAAAGCCGTGCATTGTCTCTCACATTATGGATGGAGAA	3333
Qy	939	nLeuIleGlnArgTyrIleTrpAspGlyGluSerAsnCysGluTrpSerIlePheTyrVa	959
Db	3334	CCTTATACAGCGTTATATCTGGGATGGAGAGAGCAACTGTGAGTGGAGTATATCTATGT	3393
Qy	959	lThrValLeuAlaPheThrIleuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCy	979
Db	3394	GACAGTGTTCGCTTTTACTCTTATTTGCTTAAACAGGAGGTTTCACTGGCTTTGGCATCTG	3453
Qy	979	sCysLysArgGlnLysArgThrLysIleArgLysLysThrLysTyrThrIleLeuAs	999
Db	3454	CTGCTGCANAAGACNAANAAGGACTAAATCAGGAAAAAANAAGTACACCATCTCGGA	3513
Qy	999	pAsnMetAspGluGlnGluArgMetGluLeuArgProLysTyrGlyIleLysHisArgSe	1019
Db	3514	TAACATGGATGAACAGGAAAGAATGGAACCTGAGGCCCAATATGGTATCAAGCACCGAAG	3573
Qy	1019	rThrGluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspTh	1039
Db	3574	CACAGACCAACTCCAGCTCGATGGTATTCCGAGTCTGAGTTTGACAGTGCACCGAGCAC	3633
Qy	1039	rIlePheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerIl	1059
Db	3634	ANCTTCAGCCGAGANAAGATGGNAGAGAGGGAATCCAAAGGTTTCCATGAATGGTCCAT	3693
Qy	1059	eArgAsnGlyAlaSerPheSerTyrCysSerLysAspArg	1072
Db	3694	CAGAAATGGAGCTTCCTTCAGTTATTCCTCAAGGACAGA	3733

RESULT 5

US-10-764-390-6	
Sequence 6, Application US/10764390	
Publication No. US2004021421A1	
GENERAL INFORMATION:	
APPLICANT: Agensys, Inc.	
APPLICANT: Raitano, Arthur B.	
APPLICANT: Jakobovits, Aya	
APPLICANT: Challita-Eid, Pia M.	
APPLICANT: Ge, Wangmao	
APPLICANT: Faris, Mary	
APPLICANT: Steven B. Kanner	
APPLICANT: Juan J. Perez-Villar	
TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins	
TITLE OF INVENTION: Entitled 254p1D6B Useful in Treatment and Detection of	
TITLE OF INVENTION: Cancer	
FILE REFERENCE: 51158-20081.00	

QY 210 nGlnAspProGluLeuHisTyrLeuAsnGluSerAlaSerThrProAlaProLysLeuPr 230
 Db 1341 GCAGGACCCCTGAGCTCCATTGCTGAATGAGTGGCTTCAACCCCTGCCCAAACTCCC 1400
 QY 230 oGluArgSerValLeuLeuProLeuProThrThrProSerSerGlyGluValLeuGluY 250
 Db 1401 TGAGAGAAGTGTGTGCTTCCCTTCCGCTACTCATCTTCCAGGAGGTGTTCGAGAA 1460
 QY 250 sGluLysAlaSerGlnLeuGlnGlnSerSerAsnSerSerGlyLysGluValLeuMe 270
 Db 1461 AGAAAAGGCTTCTCAGCTCCAGAACATCCAGCAACAGCTCTGAAAAGAGGTCTTAAT 1520
 QY 270 tProSerHisSerLeuProProAlaSerLeuGluLeuSerSerValThrValGluLysSe 290
 Db 1521 GCCTTCCCATAGTCTTCCCTCCGCGAAGCCTGGAGCTCAGCTCAGTCACCGTGGAGAAAAG 1580
 QY 290 rProValLeuThrValThrProGlySerThrGluHisSerIleProThrProProThrSe 310
 Db 1581 CCCAGTGTCTCACAGTCACCCCGGGAGTACAGAGCACAGCATCCCAACACCTCCCACTAG 1640
 QY 310 rAlaAlaProSerGluSerThrProSerGluLeuProIleSerProThrThrAlaProAr 330
 Db 1641 CCAGCCCCCTCTGAGTCCACCCCATCTGAGCTACCCATATCTCTTACCACTGTCTCCAG 1700
 QY 330 gThrValLysGluLeuThrValSerAlaGlyAspAsnLeuIleIleThrLeuProAspAs 350
 Db 1701 GACAGTGAAGAAGCTTACGGTATCGGCTCGAGATAACCTAAATATACTTTTACCAGCAA 1760
 QY 350 nGluValGluLeuLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTy 370
 Db 1761 TGAAGTGAACCTGAAGGCCCTTGTTCGCGCAGCGCCACTGTAGAAACAACCTACAACCTA 1820
 QY 370 rGluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGluIleLysGlnGlyHisly 390
 Db 1821 TGAATGGAATTAATAAGCCACCCACAGACTACCAAGGTGAATAAACAAGGACACAA 1880
 QY 390 sGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSe 410
 Db 1881 GCNAACCTTAACTCTCTCAATTGTCGCTCGGACTTTATGCTTCAAGTCACTGTTTC 1940
 QY 410 rSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgVa 430
 Db 1941 TAGTGAAAACCGCTTTGGAGAAGGATTTGTCAATGTCACTGTAAGCCCTGCCAGAAGAGT 2000
 QY 430 lAsnLeuProValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuTh 450
 Db 2001 CAACCTGCCACCTGTAGCAGTTGTTTTCTCCCCAAGTGCAGAGCTCAGCTTCCCTTTGAC 2060
 QY 450 rSerAlaLeuIleAspGlySerGlnSerThrAspThrThrGluIleValSerTyrHisTr 470
 Db 2061 GTCAGCCCTCATTTGATGGCAGCAAAAGTACAGATGATAGTGAATAGTCAATTATCATTTG 2120
 QY 470 pGluGluIleAsnGlyProPheIleGluGluLysThrSerValAspSerProValLeuAr 490
 Db 2121 GGAAGAAATAAAGCGGCCCTTACAGAGAAGACTTCAGTTGACTTCTCCGCTCTTAGC 2180
 QY 490 qLeuSerAsnLeuAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGl 510
 Db 2181 CTTGTCTAACCTTGATCTCGTGAATATATAGTTTTCAGGTTGACTGTTTACAGACTCGGACGG 2240
 QY 510 yAlaThrAsnSerThrAlaAlaLeuIleValAsnAsnAlaValAspTyrProProVa 530
 Db 2241 AGCCACTAACTCTACAACTGCAGCCCTAATAGTGAACAATGCTGTGGACTACCCACACT 2300
 QY 530 lAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGl 550
 Db 2301 TGTAAATGCAAGGACCAAAATCACACCACTTATGCCCCCAAACTCCATCACTTTGAATGG 2360
 QY 550 yAsnGlnSerSerAspAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySe 570
 Db 2361 AAACCAAGAGTACAGATCACAGATTTGCTCTATAGTGTGCTCCCTGGGCTCGGGAG 2420
 QY 570 rGluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMe 590

Db 2421 TGAGGGCAAAACATGTGTGTCATGCGGGAGTACAGACGCCATACCTTCAATTTTCTGCAAT 2480
 QY 590 tGlnGluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerTh 610
 Db 2481 GCAGAAAGGAGATATATACATTTTCAGCTGAAAGGTGACAGATTTCTTCAAGGCAACAGTCTAC 2540
 QY 610 rAla***ValThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaGl 630
 Db 2541 TGCTGTGTGACTGTGATTTGCCAGCTTGAACCAATAGACCTCCAGTGGCTGTGGCCGG 2600
 QY 630 yProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSe 650
 Db 2601 CCCTGATAAAGAGCTGATCTTCCAGTGGAAAGTGTACCTCGATGGATGGAGCAGCAG 2660
 QY 650 rAspAspHisGlyIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValGl 670
 Db 2661 CGATGACCACGGCATTTCTTCTACCTGGGAGCAGCTCAGAGGCCCTCAGTGCAGTGA 2720
 QY 670 uMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHi 690
 Db 2721 GATGNAATAATTGACAAAGCAATAGCCACTGTGACTGTCTCCAGGTGGGGACCTACCA 2780
 QY 690 sPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAl 710
 Db 2781 CTTCCGTTTGACAGTGAAGAACCCAGCAGGGACTGAGCAGCAGCTCCACCTCACTGTGGC 2840
 QY 710 aValLysLysGluAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuVa 730
 Db 2841 TGTGAAGAAGAAATAATAGTCTCTCCAGAGCCCGGCTGGTGACACATGTTCTTGT 2900
 QY 730 lLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleVa 750
 Db 2901 GCTTCCCAATATTTCCATTACTTTGGATGGTTCAAGGTCTACTGATGACCAAGAATTGT 2960
 QY 750 lSerTyrLeuTrpIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySe 770
 Db 2961 GTCTCTCTGTGGATCGGGATGGCCAGAGTCCAGCAGCTGGAGATGTCATCGATGGCTC 3020
 QY 770 rAspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisIle 790
 Db 3021 TGACCACAGTGTGGCTTGCAGCTTACGAATCTGTGTGAGGGGGTGTACACTTTCCACTT 3080
 QY 790 uArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnPr 810
 Db 3081 GCGAGTCAACCCAGCTCAGGGGGCTCGGACACAGACACTGCCACTGTGGAAGTCAGCC 3140
 QY 810 oAspProArgLysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuTh 830
 Db 3141 AGACCCTAGGAAGAGTGGCTGGTGGAGTGCACCTGCAGGTTGGTGTGGGCAGCTGAC 3200
 QY 830 rGluGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSe 850
 Db 3201 AGAGCAGCGGAAGGACACCTTTGTGAGGAGCAGCTGGCTGTGCTGGAACGCTGGACTC 3260
 QY 850 rAspIleLysValGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTy 870
 Db 3261 GGACATTAAGGTCCAGAAGATTCGGGCCCACTCGGATCTCAGCACCGTATTGTGTTTA 3320
 QY 870 rValGlnSerArgProProPheLysValLeuLysAlaAlaGluValAlaArgAsnLeuHi 890
 Db 3321 TGTCAGAGCAGCGCCCTTTCAAGGTTCTCAAGCTGCTGAAGTGGGCCGAAATCTGCA 3380
 QY 890 sMetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAspTh 910
 Db 3381 CATCGGCTCTCAAGAGGAGGCTGACTTCTTGTCTTTTCAAGGTTCTGAGGGTTGATAC 3440
 QY 910 rAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCy 930
 Db 3441 AGCAGTTGCCCTTCTGAAGTGTTCCTGGCCATGTTCACTGCGACCCCTCACAAAGCGCTG 3500
 QY 930 sIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyGluSe 950

Db 3501 CATTGCTCTCACTTATGGATGGAGAACCTTATACAGCGTTATATCTCGGATGGAGAG 3560
 QY 950 rAsnCysGluTrpSerIlePheTyrValThrValLeuAlaPheThrLeuIleValLeuTh 970
 Db 3561 CAACGTGAGTGGAGTATATTCTATGTGACAGTGTGGCTTTTACTCTTATTGTGCTAAC 3620
 QY 970 rGlyGlyPheThrTrpLeuCysIleCysCysCysIleArgGlnLysValArgThrLysIleAr 990
 Db 3621 AGGAGTTTCACTTGGCTTTGCACTCTGCTGCAAAAGACAAAAGAGACTAAATCAG 3680
 QY 990 gLysLysThrLysTyrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuAr 1010
 Db 3681 GAAAAACAAAGTACACCATCTGGATAACATGGATCAACAGGAAAGAAATGGAACCTGAG 3740
 QY 1010 gProLysTyrGlyLysIleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerG1 1030
 Db 3741 GCCCAATATGTTATCAAGCACCCGAGCACAGAGCACAACTCCAGCTGTATGATCCGA 3800
 QY 1030 uSerGluPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAs 1050
 Db 3801 GTCTGAGTTTGACAGTGAACAGGACACAAATCTTCAGCCGAGAAAGATGGAGAGGGAA 3860
 QY 1050 nProLysValSerMetAsnGlySerIleArgAsnGlyAlaSerPheSerTyrCysSerLy 1070
 Db 3861 TCCAAAGGTTTCCATGATGTTCCATCAGAAATGGAGCTTCCTTCAGTTATTGCTCAA 3920
 QY 1070 sAspArg 1072
 Db 3921 GGACAGA 3927

RESULT 6

US-10-764-390-269

; Sequence 269, Application US/10764390

; Publication No. US20040214212A1

; GENERAL INFORMATION:

; APPLICANT: Agensys, Inc.

; APPLICANT: Raitano, Arthur B.

; APPLICANT: Jakobovits, Aya

; APPLICANT: Challita-Eid, Pia M.

; APPLICANT: Ge, Wangmao

; APPLICANT: Faris, Mary

; APPLICANT: Steven B. Kanner

; APPLICANT: Juan J. Perez-Villar

; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins

; TITLE OF INVENTION: Entitled 2541D6B Useful in Treatment and Detection of

; FILE REFERENCE: 51158-20081.00

; CURRENT APPLICATION NUMBER: US/10/764,390

; PRIOR FILING DATE: 2004-01-23

; PRIOR FILING DATE: 2003-01-24

; NUMBER OF SEQ ID NOS: 277

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 269

; LENGTH: 6991

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-764-390-269

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0	5527.00	6991	1068
Percent Similarity:	96.83%	Conservative:	0
Best Local Similarity:	96.83%	Mismatches:	4
Query Match:	99.05%	Indels:	31
DB:	20	Gaps:	1

US-10-764-390-3 (1-1072) x US-10-764-390-269 (1-6991)

QY 1 MetAlaProThrGlyValLeuSerSerLeuLeuLeuValThrIleAla----- 18

Db 621 ATGGGCCCCCAGAGGTGCTCTCTTCATTGCTGCTGCTGGTGCACAAATTGCAGTTTGC 680
 QY 350 nGluValGluLeuLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTy 370

QY 18 ----- 18
 Db 681 TTATGGTGGATGCACATCATGGCAAAAAAATCATCGTGGAGCATCATTTAAGAAGACCCAT 740
 QY 19 -----GlyCyAlaArgLysGlnCysSerGluGlyArgTh 30
 Db 741 GACTAGACTGGCTGGCCGAGCCCATGTTGTGCCGTAAAGCAGTGCAGCGAGGGAGGAC 800
 QY 30 rTyrSerAsnAlaValIleSerProAsnLeuGluThrThrArgIleMetArgValSerHi 50
 Db 801 ATATTCCAATCAGCTATTTTCACTTAATCTGGAAAAACACAGAAATCATGCGGGTGTCTCA 860
 QY 50 sThrPheProValValAspCysThrAlaAlaCysCysAspLeuSerSerCysAspLeuAl 70
 Db 861 CACCTTCCTCTGCTGAGACTGCAGCGCGCTGTGTGACCTGTCCAGCTGTGACCTGGC 920
 QY 70 aTriPheGluGlyArgCysTyrLeuValSerCysProHisLysGluAsnCysGluPr 90
 Db 921 CTGGTGGTTTCAGGGCGCGCTGCTACCTGGTGGAGTGGCCCCCAAGAGAACTGTGAGCC 980
 QY 90 oLysLysMetGlyProIleArgSerTyrLeuThrPheValLeuArgProValGlnArgPr 110
 Db 981 CAAGAAGATGGCCCATCAGGTCTTATCTCCTTTGTGCTCCGGCTGTTCAGAGGCC 1040
 QY 110 oAlaGlnLeuLeuAspTyrGlyAspMetMetLeuAsnArgGlySerProSerGlyIleTr 130
 Db 1041 TGCACAGCTGCTGACTATGGGGACATGATGCTGAACAGGGGCTCCCCCTCGGGATCTG 1100
 QY 130 pGlyAspSerProGluAspIleArgLysAspLeu***PheLeuGlyLysAspTriGlyLe 150
 Db 1101 GGGGACTCCTCCTGAGGATATCAGAAAGACTTTCCTTCTAGGCAAGATTGGGGCT 1160
 QY 150 uGluGluMetSerGluTyr***AspAspTyrArgGluLeuGluLysAspLeuLeuGlnPr 170
 Db 1161 AGAGGAGATGCTGAGTACTCAGATGACTACCGGGAGCTGGAGAAGGACCTCTTCAAC 1220
 QY 170 oSerGlyLysGlnGluProArgGlySerAlaGluTyrThrAspTrpGlyLeuLeuProGl 190
 Db 1221 CAGTGGCAAGCAGAGGCCCGAGAGGAGTGGCGAGTACACGAGCTGGGGCTTACTGCCGG 1280
 QY 190 ySerGluGlyAlaPheAsnSerSerValGlyAspSerProAlaValProAlaGluThrGl 210
 Db 1281 CAGCAGGGGGCTTCACTCTCTGTTGGAGACAGTCTCGGGTGCACGGAGACCGCA 1340
 QY 210 nGlnAspProGluLeuHisTyrLeuAsnGlnSerAlaSerThrProAlaProLysLeuPr 230
 Db 1341 GCAGGACCTCGAGCTCCATTACCTGAATGAGTGGCTTCAACCCCTGCCCAAACTCCC 1400
 QY 230 oGluArgSerValLeuLeuProLeuProThrThrProSerSerGlyLysValLeuGluY 250
 Db 1401 TGAGAGAAAGTGTGTGCTTCCCTTGGCGACTACTCCATCTTCAGGAGAGGTGTGGAGAA 1460
 QY 250 sGluLysAlaSerGlnLeuGlnGlnSerSerAsnSerSerGlyLysGluValLeuMe 270
 Db 1461 AGAAAAGCTTCTCAGCTCCAGGAACAATCCAGCAACAGCTCTGGAAAAGAGGTTCAT 1520
 QY 270 tProSerHisSerLeuProAlaSerLeuLeuLeuSerSerValThrValGluLysSe 290
 Db 1521 GCCTTCCCATAGTCTTCTCCCGCAAGCTCGAGCTCAGCTCAGTCCCGTGGAGAAAAG 1580
 QY 290 rProValLeuThrValThrProGlySerThrGluHisSerIleProThrProThrSe 310
 Db 1581 CCCAGTGTCTACAGTACCCCGGGAGGTACAGAGCAGACATCCCAACACCTCCACCTAG 1640
 QY 310 rAlaAlaProSerGluSerThrProSerGluLeuProIleSerProThrThrAlaProAr 330
 Db 1641 CGCAGCCCCCTCTCAGTCCACCCCATCTGAGCTACCCATATCTCTACCTGCTCCAG 1700
 QY 330 gThrValLysGluLeuThrValSerAlaGlyAspAsnLeuIleIleThrLeuProAspAs 350
 Db 1701 GACAGTGAAGAAGACTTACGGTATCGGCTGGAGATACTTAATATATACTTATCCCGACA 1760
 QY 350 nGluValGluLeuLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTy 370

Db 1761 |||||TGAAGTTGAACCTGAAGGCGCTTTGTTGCGCGAGCGCCACCTGTAGAAAAACAACCTACAACTA 1820
 Qy 370 rGluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGluIleLysGlnGlyHis 390
 Db 1821 TGAATGGAAATTAATAAGCCACCCACAGACTACCAAGGTGAATATAAACCAAGGACACAA 1880
 Qy 390 sGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSe 410
 Db 1881 GCRAACTCTTAACCTCTCTCAATTTGTCGTGCGACTTTATGTTCTTCAAGTCACTGTTTC 1940
 Qy 410 rSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgVa 430
 Db 1941 TAGTGAAAAACGCGCTTTGGAGAAAGGATTTGTCAATGTCACTGTTAAGCCTGCCAGAAGAGT 2000
 Qy 430 lAsnLeuProProValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuTh 450
 Db 2001 CAACTTGCCACCTGTAGCAGTTGTTTCTCCCACTGCAAGAGCTCACTTTGCCCTTTGAC 2060
 Qy 450 rSerAlaLeuIleAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHisTr 470
 Db 2061 GTGAGCCCTCATGTGTCGAGCCCAAGTACAGATGATGAAATAGTAGTATCATTTG 2120
 Qy 470 pGluIleAsnGlyProPheIleGluLysThrSerValAspSerProValLeuAr 490
 Db 2121 GGAAGAAATAAACGGGCGCTTCATAGAGAGAAAGACTTCAGTTGACTCTCCCGTCTTACG 2180
 Qy 490 gLeuSerAsnLeuAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGl 510
 Db 2181 CTTGTCTAACTTGATCTCCGGTAACTATATAGTTTTCAGGTTGACTGTTTACAGACTCGGACGG 2240
 Qy 510 yAlaThrAsnSerThrAlaAlaLeuIleValAsnAsnAlaValAspTyrProProVa 530
 Db 2241 AGCCACTAACTCTACAACTGCGAGCCCTAATAGTGAACATGCTGTGGACTACCCACAGT 2300
 Qy 530 lAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGl 550
 Db 2301 TGCTAATGCGAGGACCAATACACACATAACTTTGGCCCCAAAACCTCCATCACTTGAATGG 2360
 Qy 550 yAsnGlnSerSerAspAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySe 570
 Db 2361 AAACCCAGACGTCAGATCACAGATTGCTCTATGATGGTCCCTGGGCTCTGGGAG 2420
 Qy 570 rGluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMe 590
 Db 2421 TGAGGCAAAACATGTGGTTCATGCGGGAGTACAGCGCCATACCTTCATTTATCTGCAAT 2480
 Qy 590 tGlnGluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerTh 610
 Db 2481 GCAGGAAGGAGATTATACATTTTCAGCTGAAGGTGACAGATTCTTCAAGGCAACAGTCTAC 2540
 Qy 610 rAla***ValThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaGl 630
 Db 2541 TGCTGTGGTGACTGTGATTTGTCAGGCTGAAAAACAATAGACCTCCAGTGGCTGTGGCCGG 2600
 Qy 630 yProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSe 650
 Db 2601 CCTGTATTAAGAGCTGATCTTCCAGTGGAAAGTGTCTACCTGGATGGAGGAGCAGCAG 2660
 Qy 650 rAspAspHisGlyIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValGl 670
 Db 2661 CGATGACCGCGCAATTTCTTACCACTGGGAGCAGCTCAGAGGCGCCCGTGCAGTGA 2720
 Qy 670 uMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHi 690
 Db 2721 GATGGAAAAATATTGACAAAGCAATAGCCATGTGACTGGTCTCCAGGTGGGGACCTTACCA 2780
 Qy 690 sPheArgLeuThrValIleAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAl 710
 Db 2781 CTTCCGTTTGACAGTGAAGACCAGCAGGAGCTGAGCAGCAGCTCCACCCTCACTGTGGC 2840
 Qy 710 aValLysLysGluAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuVa 730

Db 2841 TGTGAAGAAGAAATAATAGTCTCCAGAGCCCGGGCTGGTGGCAGACATGTTCTTGT 2900
 Qy 730 lLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleVa 750
 Db 2901 CTTTCCCAATTAATTTCCATTACTTTGGATGGTTTCAAGGTCTACTGATGACCAAGATTTGT 2960
 Qy 750 lSerTyrLeuTrpIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySe 770
 Db 2961 GTCTCTATCTGTGGATCCGGGATGGCCAGAGTCCAGAGCTCGAGATGTCATCGATGGCTC 3020
 Qy 770 rAspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisIle 790
 Db 3021 TGACCACAGTGTGGCTCTGCAGCTTACGAATCTGTGTGAGGGGGTGTACACTTTTCCACTT 3080
 Qy 790 uArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnPr 810
 Db 3081 CCGAGTCCACCACAGTCCAGGGGGCTCGEACACAGACACTGCCACTGTGGAGTGCAGCC 3140
 Qy 810 oAspProArgLysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuTh 830
 Db 3141 AGACCCTAGGAAGAGTGGCTGTGGAGCTGACCTGCAGGTGGTGTGGCAGCTGAC 3200
 Qy 830 rGluGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSe 850
 Db 3201 AGAGCAGCGAAGGACACCCCTTGTGAGCAGCTGGCTGTGCTGGAACGTGCTGGACTC 3260
 Qy 850 rAspIleLysValGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTy 870
 Db 3261 GGACATTAAGGTCCAGAGATTCGGGCCCACTTCGGATCTCAGCACCGTGATGTGTTTA 3320
 Qy 870 rValGlnSerArgProPropheLysValLeuLysAlaAlaGluValAlaAlaArgAsnLeuHi 890
 Db 3321 TGTACAGAGCAGGCGCGCTTCAAGGTTCTCAAAGCTGCTGAAGTGGCCGGAATCTGCA 3380
 Qy 890 sMetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAspTh 910
 Db 3381 CATCGGCTCTCAAAGGAGAGGCTGACTTCTTGTCTTTTCAAGGCTTTGAGGGTGTGATAC 3440
 Qy 910 rAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCy 930
 Db 3441 AGCAGGTTGCTTCTGAAGTGTCTGGCCATGGTCACTCGCACCCCTCACAAGGCGCTG 3500
 Qy 930 sIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyLys 950
 Db 3501 CATTTGCTCTCATTTATGATGGAGAACCTTATACAGCGTTATATCTGGGATGGAGAGAG 3560
 Qy 950 rAsnCysGluTrpSerIlePheTyrValThrValLeuAlaPheThrLeuIleValLeuTh 970
 Db 3561 CAACTGTGAGTGGAGTATATTTCTATGTGACAGTGTGGCTTTTACTCTTTATTTGTGCTAAC 3620
 Qy 970 rGlyGlyPheThrTrpLeuCysIleCysCysLysArgGlnLysArgThrLysIleAr 990
 Db 3621 AGGAGGTTTCACTGGCTTTGCACTGTCTGTGCAAAAGACAAAAAGAGACTAAAAATCAG 3680
 Qy 990 glyLysThrLysTyrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuAr 1010
 Db 3681 GAAAAAACAAGTAGTACCATCTCGATAACATGATGAACAGGAAAGAAAGTGAAGTGAAG 3740
 Qy 1010 gProLysTyrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGl 1030
 Db 3741 GCCCAAAATATGGTATCAAGCACCGAAGCACAGACCAACTCCAGCTGATGGTATCCGA 3800
 Qy 1030 uSerGluPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAs 1050
 Db 3801 GTCTGAGTTTGACGTGACGAGCACCAATCTTTCAGCCGAGAAAGATGAGAGAGGGAA 3860
 Qy 1050 nProLysValSerMetAsnGlySerIleArgAsnGlyAlaSerPheSerTyrCysSerLy 1070
 Db 3861 TCCAAAAGGTTTCCATGAATGGTTTCCATCAGAAATGGAGCTTCTTCAAGTTATTTGCTCAAA 3920
 Qy 1070 sAspArg 1072
 Db 3921 GGACAGA 3927

RESULT 7
 US-11-097-143-16883
 ; Sequence 16883, Application US/11097143
 ; Publication No. US20050208558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: et al
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 ; TITLE OF INVENTION: DROSOPHILA GENES.
 ; FILE REFERENCE: CL000728
 ; CURRENT APPLICATION NUMBER: US/11/097,143
 ; PRIOR FILING DATE: 2005-04-04
 ; PRIOR APPLICATION NUMBER: 60/157,832
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: 60/160,191
 ; PRIOR FILING DATE: 1999-10-19
 ; PRIOR APPLICATION NUMBER: 60/161,932
 ; PRIOR FILING DATE: 1999-10-28
 ; PRIOR APPLICATION NUMBER: 60/164,769
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: 60/173,383
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: 60/175,693
 ; PRIOR FILING DATE: 2000-01-12
 ; PRIOR APPLICATION NUMBER: 60/184,831
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/191,637
 ; PRIOR FILING DATE: 2000-03-23
 ; NUMBER OF SEQ ID NOS: 43008
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16883
 ; LENGTH: 3345
 ; TYPE: DNA
 ; ORGANISM: DROSOPHILA
 US-11-097-143-16883

Alignment Scores:
 Pred. No.: 1.08e-121 Length: 3345
 Score: 1451.50 Matches: 373
 Percent Similarity: 46.94% Conservative: 172
 Best Local Similarity: 32.13% Mismatches: 383
 Query Match: 26.01% Indels: 233
 DB: 24 Gaps: 35

US-10-764-390-3 (1-1072) x US-11-097-143-16883 (1-3345)

Qy	8	LeuSerSerLeuLeuLeuValThrIleAlaGlyCysAla-----	21
Db	157	ATTTGCAACCTGCTTTTGTGCTGCCACTGCGATGTCGTCAGCATATGCCGAGCTCACAAAG	216
Qy	22	-----ArgLysGlnCysSerGluGlyArgThrTyr 31	
Db	217	CAAAACGCTGTGTTGGTGGAGTAAGAAACACAGGAGACCACCCCGATAATTCAGTT	276
Qy	32	SerAsnAlaValIleSerProAsnLeuGluThrThrArgIleMetArgValSerHisThr 51	
Db	277	GGCGGTTCT---ATATCGCCCAATTTGTGTGCGCCACAGATGCTGAGG-----CATGTG 327	
Qy	52	Pro 53	
Db	328	TTTCAAAACGCCACTCCCGGGGATGAGCAGCAGCGGGGTGTGTTGAGGAGTACAAGCCA 387	
Qy	54	ValValAsp-----CysThrAlaAla 60	
Db	388	CCGCGGATGAGTGGAGCCACTGGAGGAGAGCGCTATCTTTGGAATGCTGTACAGGCT 447	
Qy	61	CysCysAsp-----LeuSerSerCysAspLeuAlaTTrpTrpPheGluGly 75	
Db	448	TGCTGCGAAAGCCCCCGAATGGCAGCAGTGCCTGCAATGTGGTCTCTGTTTAAAGCC 507	
Qy	76	ArgCysTyrLeuValSerCysProHisLysGluAsnCysGluProLys-----Lys 92	

Db	508	AAATGCTACCATATCCGATGCCAGAGCAACGAGCCCTGCTTGTCCCAAGCTCAGAGTTCCG 567
Qy	93	MetGlyProIleArgSerTyrLeuThrPheValLeuArgProVal----- 107
Db	568	ATG---CCCAACGAAAGGTTTCAGATGGTACTGCTCAATCCACTGGCGGATGCCACATGG 624
Qy	108	-----GlnArgProAlaGlnLeuLeuLeuAspTyrGlyAsp 118
Db	625	CCACAGCTCTCAAGCGGAGCGACCAACAGAAATGCTGAGATTCTTCCATACGATGAA 684
Qy	119	MetMetLeuAsnArgGlySerProSerGlyIleTyrGlyAspSerProGluAspIleArg 138
Db	685	CGCGCACTTAAT-----TTCTGGAAGCAGCCG----- 711
Qy	139	LysAspLeu***PheLeuGlyLysAspTrpGlyLeuGluGluMetSerGluTyr***Asp 158
Db	712	AGAAAGTTGAGCTACTGCTCGCAAT-----CAGGAACCCCGTTTATGAGGAC 762
Qy	159	Asp-----TyrArgGluLeuGlyLysAspLeuLeuGlnProSerGlyLys 173
Db	763	GAAGATTTCCTTGGCGGACAGCAATGAATCAATGATCTTCCACCCGCGAAGAAC 822
Qy	174	GlnGluProArgGlySerAlaGluTyrThrAspTrpGlyLeuLeuProGlySerGluGly 193
Db	822	----- 822
Qy	194	AlaPheAsnSerSerValGlyAspSerProAlaValProAlaGluThrGlnGlnAspPro 213
Db	823	-----GATGTTCTGCCCAACGAA 840
Qy	214	GluLeuHisTyrLeuAsnGluSerAla-----SerThrProAla 226
Db	841	GAGCTCGGCTACTACGATTCGAATGCCAAGTTTACCACCTGCGATATGGAGACACCTTGT 900
Qy	227	ProLysLeuProGluArgSerValLeuLeuProLeuProThrThrProSerSer----- 244
Db	901	CCACCA---CCGCAACCAATGTGTA-----CCACTACAGCCAAATGCCGTGCGT 945
Qy	245	-----GlyGluValLeuGlyLysGluLys----- 252
Db	946	GGAGTGTGCACCTGCTCCAGAAAGGCTTTGTGGGAATAAACAAGAGAGTGGTAATGGCA 1005
Qy	253	-----AlaSerGlnLeuGlnGluSerSerAsnSerSerGlyLysGluVal 268
Db	1006	CGGTTCTCCCTATAGCTCATATCTA-----ACCAGCAATGAGCGGGACAGCAAGAA 1056
Qy	269	LeuMetProSerHisSerLeuProProAlaSerLeuGluLeuSerSerValThrValGlu 288
Db	1057	CGGCTGCTAGTAGAAGTCACTCCCGAGGTATCA----- 1089
Qy	289	LysSerProValLeuThrValThrProGlySerThrGluHisSerIleProThrProPro 308
Db	1090	-----ACGCGGCT 1098
Qy	309	ThrSerAlaAlaProSerGluSerThrThrProSerGluLeuProIleSerProThrThra 328
Db	1099	TTGAAGCTGAGCAGAAC----- 1116
Qy	329	ProArgThrValLysGluLeuThrValSerAlaGlyAspAsnLeuIleIleThrIleuPro 348
Db	1117	-----AAGGACATTGTAGTCTCGTAAATGTCGAAG---GAAGTTGTTGTTGCGG 1161
Qy	349	AspAsnGluValGluLeuLysAlaPheValAlaProAlaProVal---GluThrThr 367
Db	1162	GAACAGAGGTGACCTTAGCGGCTTTACGTTCCCGATGAGCAACACAGCGATACCAAG 1221
Qy	368	TyrAsnTyrGluTrpAsnLeuIleSerHisProThr---AspTyrGlnGlyGluIleLys 386
Db	1222	TACAAATATCTTTGGACACTTATATCGCAACCCCAAGGGTCTCATGAATGGCACTATTTCG 1281
Qy	387	GlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLys 406

Db 1282 GATCAGACAAAGTCGAAGGTAAACTATCGAATCTTTTCGGAGGAGCTGTATATCTTTTAAG 1341
 Qy ValThrValSerSerGluAsn---AlaPheGlyGluGlyPheValAsnValThrVallys 425
 Db 1342 GTCACTGTAACTGTGACAAATGGAAACCTTTGGCGAGGCAACAGCCAATGTTACAGTGT 1401
 Qy 426 ProAlaArgValAsnLeuProProValAlaValValSerProGlnLeuGlnGluLeu 445
 Db 1402 CCGGAGAATCGAATCAATCAGCCACCAACAGTCATCATCTCGCCCGAGGAGCAGATCATC 1461
 Qy 446 ThrLeuProLeuThrSerAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 465
 Db 1462 CGTCAGCCACCAACCAATGCTTATGATGCGCAGCAGCAGTACGATGATGATGATGAT 1521
 Qy 466 ValSerTyrHisTrpGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 485
 Db 1522 ACCAATGGCACTGGGAGGTGATTTCCGGACCGATGTTGTTACCAACCAAGTTTTCGCGAG 1581
 Qy 486 SerProValLeuArgLeuSerAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 505
 Db 1582 GTTAACACCTTCAGTTGGATCTCATATCGCCGGGGAACACATTCAGAGTAAACCGTA 1641
 Qy 506 ThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuLeuLeuLeuLeuLeu 525
 Db 1642 ACCGATTCAAAATACGTCACCAATTCACCACTGCGCAGATAGCAGTCTTTAAGGAGACC 1701
 Qy 526 AspTyrProProValAlaAsnAlaGlyProAsnHisThrThrLeuLeuProGlnAsnSer 545
 Db 1702 GATTATGCTCGGTGGCAATCCCGCGAGTCCGATGATCTTGATTTGCGCAACCAATAT 1761
 Qy 546 IleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleValLeuLeuLeuLeuLeu 565
 Db 1762 GTCACTCGAATGGCAGACCCAGTTCGGATGATCAGAGATGTTGCTGGGAGTGGACC 1821
 Qy 566 LeuGlyProGlySerGluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeu 585
 Db 1822 AAGGATGCTAGCGACGCAAGCTGTGGATATGCAAGATACAGAAACACCTATGTT 1881
 Qy 586 HisLeuSerAlaMetGlnGlyAspTyrThrPheGlnLeuLeuValThrAspSerSer 605
 Db 1882 CAGCTGTCCAATTTGGAGAGGCGCATGTACATTTTGTCTAAAGTAAACCGATGGCAGT 1941
 Qy 606 ArgGlnGlnSerThrAla***ValThrValIleValGlnProGluAsnAsnArgProPro 625
 Db 1942 GGGCAATCGAGTACAGTAAAGTCCATGTTTGTGAAGCTCCCAACCAATTCACCA 2001
 Qy 626 ValAlaValAlaGlyProAspLysGluLeuIlePheProValGluSerAlaThrLeuAsp 645
 Db 2002 GTTGCTGAGCGAGGAAGTAAATACGACTACTAGCTTGCCCATTAATTTGGGTTCTTTTGAAT 2061
 Qy 646 GlySerSerSerSerAspAspHisGlyIleValPheTyrHisTrpGluHisValArgGly 665
 Db 2062 GGCTCCGATTCAAGAGACACATTTGGCATCAAGAGTTATTTGTGAAGCAGCTGAGCGGG 2121
 Qy 666 ProSerAlaValGluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGln 685
 Db 2122 CCCAATAATGCACTCAATTTGAAGTCCAACTCATCAATGCAATGCCACCTCTCTGACT 2181
 Qy 686 ValGlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrSer 705
 Db 2182 CTGGGACTCTATGAGTTCCGAATTAACCGTAGCTGATGAAATAATAAACAACCGCTACGGAC 2241
 Qy 706 ThrLeuThrValAlaValLysLysGluAsnAsnSerProProArgAlaAlaGlyGly 725
 Db 2242 ACCACGTGGTGAAGATAGTTCAAGAACCGAACCGACGCTCCCAATAGCCCAACCGGTGT 2301
 Qy 726 ArgHisValLeuValLeuProAsnSerIleThrLeuAspGlySerArgSerThrAsp 745
 Db 2302 GATCACACCGTCACTCCGCGGCCCGCCATCTATTTCAATGGCTCTTAATCTCTGGAT 2361
 Qy 746 AspGlnArgIleValSerTyrLeuTrpIleArgAspGlyGlnSerProAlaAlaGlyAsp 765
 Db 2362 GACCTGGCTGTGTCAGATATCTCTGGACACGCGACGAGCAGATTTTGGCAGCGGCGCTC 2421

Qy 766 ValIleAspGlySerAspHisSerValAlaLeuLeuGlnLeuThrAsnLeuValGluGlyVal 785
 Db 2422 ATTGTGGCAGATACCGACAGGAGCCCGTAATGATTTTGACCAATTTAGTACAGCGCGC 2481
 Qy 786 TyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThr 805
 Db 2482 TATGTATTCACTGACTGTGAGCGATGATCAGGGTTTGACCAAGTTTCAGATCTGTACG 2541
 Qy 806 ValGluValGlnProAspProArgLysSerGlyLeuValGluLeuLeuGlnValGly 825
 Db 2542 GTTAATGTCCGTCGCGATCCCAAGCTATGTAATTTGGTTCAATGACCCCTGCCATGGCC 2601
 Qy 826 ValGlyLeuLeuThrGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuLeuLeu 845
 Db 2602 ATCTCTGTCTCGTCAATCCGAATCGATCGGTGGTGCAAAACTGCAAGTTGTGCTG 2661
 Qy 846 AsnValLeuAspSerAspIleLysValGlnLysIleArg-----AlaHisSerAspLeu 863
 Db 2662 GGAGAT---GAGAATAAGATCCAGTCAGGAGCTAAATATATGACTTCATACGAT--- 2715
 Qy 864 SerThrValIleValPheTyrValGlnSerArgProProPheLysValLeuLysAlaAla 883
 Db 2716 GCTACTGTCTGCTGTTCTTCTAGTAAACGATCGACAGGGT---AAGGCATTTGATGGTTG 2772
 Qy 884 GluValAlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAspPheLeuLeuPhe 903
 Db 2773 CAAGTGGAGCGCAGTTGAGAACCACTCAGAAAGATGATCTATACTGGCGCATTT 2832
 Qy 904 LysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCys 923
 Db 2833 GCAGTG---GATATTGCGCACCTCTGTTGTCAGAGTGAATGCTCCGGTCATGGGTCGT 2889
 Qy 924 AspProLeuThrLysArgCysIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArg 943
 Db 2890 AATCCCATTCACAGAGCTTGATATGCGAGCGCTTCGGATGCCATCG-----GCTGGC 2943
 Qy 944 TyrIleTrpAspGly---GluSerAsnCysGluTrpSerIlePheTyrValThrValLeu 962
 Db 2944 TACTTCTTATTAACAGGAGCGCCAAATGCGAGTTCATATATATATATATATATATAT 3003
 Qy 963 AlaPheThrLeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCysCysLys 982
 Db 3004 GTGATTGGGGCTGCTCTCTGCTATCTGAGTATTTTGGGGCATAGCCTGCTGCTGCAGA 3063
 Qy 983 ArgGlnLysArgThrLysIleArgLysLysThr---LysTyrThrIleLeuAspAsnMet 1001
 Db 3064 CAATCGAAAAAGCCCGCTTCTCGTCAAGAAAGTGCAAAATATTCATTGATTGTAATAG 3123
 Qy 1002 AspGluGlnGluArgMetGluLeuArgProLysTyrGlyIleLysHisArgSerThrGlu 1021
 Db 3124 GATGAGAGCAGCC----- 3138
 Qy 1022 HisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIlePhe 1041
 Db 3139 ---AATATTCTCGAAACACATCGCTACGGAATCAGAGCGAATCCGATCTCTCTTTT 3195
 Qy 1042 SerArgGluLysMetGluArgLysAsnProLysValSerMetAsnGlnSerIleArgAsn 1061
 Db 3196 GAGACCCCGACCAATGCAATGGTCTGGGCAAGCAAGTCCGACCAATTTCTCAGATCAT 3255
 Qy 1062 Gly 1062
 Db 3256 GGA 3258

RESULT 8

US-10-450-763-8985
 ; Sequence 8985, Application US/10450763
 ; Publication No. US20050196754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIF3/US

; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 8985
 ; LENGTH: 2423
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIMILAR
 ; LOCATION: (479)..(1723)
 ; OTHER INFORMATION: 60% homologous to Homo sapiens dJ73M23.3 (KIAA0319), accession
 ; OTHER INFORMATION: number AL031230, Smith-Waterman Score=1025.
 US-10-450-763-8985

Alignment Scores:
 Pred. No.: 2,52e-109 Length: 2423
 Score: 1314.00 Matches: 358
 Percent Similarity: 53.20% Conservative: 41
 Best Local Similarity: 47.73% Mismatches: 164
 Query Match: 23.55% Indels: 188
 DB: 22 Gaps: 23

US-10-764-390-3 (1-1072) x US-10-450-763-8985 (1-2423)

Qy	365	GlutThrThrTyrAsnTyrGluThrAsnLeuIleSerHisProThrAspTyrGlnGlyGlu	384	458	TTTTGATTGCT	-----ATTACTTAGGGAGTACAG	487
Db	62	GAACAACCTACAACTATGAATGGAATTTAATAGCCTCTACGTAGTGGACCCATCAGT	121	582	ThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAspTyrThrPheGlnLeuIleVal	601	
Qy	385	IleIysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrVal	404	488	ACGCATACCTTCATTATCTGCAATGCAGGAAGAGATTATACATTTTCAGCTCAGGA	547	
Db	122	CTTGATGTTGAACCTGACAACTTTCTTTATGTTACAGTTGTCCGTGGACTTTATGTC	181	602	ThrAspSerSerArgGlnGlnSerThrAla**ValThrValIleValGlnProGluAsn	621	
Qy	405	PheIysValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrVal	424	548	TTGCAGACTTAATGCAAGAGAGATAATCTTTTTCACACTGTTTCTGTAAATCCAGAAAC	607	
Db	182	TTCAATGTG-----GGTAAGACATATGCCCTACACAGATGGGCTCAGTCACTGTC	232	622	AsnArgProProValAlaValAlaGlyProAspLysGlnLeuIlePheProValGluSer	641	
Qy	425	-----LysProAlaArgValAsnLeuProProValAlaValValSerProGln	441	608	AATAGACCTCAGCTGGCTGTGGCCGCTGATAAGAGAGCTGATCTTCCAGTGGAAAGT	667	
Db	233	CTTTTTTTTGTACAGCCAGAGAGTCAACCTGCCACCTGTAGCAGTTTATAAACTTATG	292	642	AlaThrLeuAspGlySerSerSerSerAspAspHisGly	654	
Qy	442	LeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThrAsp	461	668	GCTACCTGC-----CTGTCATATTCAAAACCCCATGCCCTCCCTTATGTTAACCACT	721	
Db	293	TCCTCCTCTCTGTTATTTCTTCTTCTTTATTTCTTAATCTTTGTAGAAAGTACAGAT	352	655	IleValPheTyrHisTyrGluHisValArgGlyProSerAlaValGluMetGluAsnIle	674	
Qy	462	AspThrGluIleValSerTyrHisTyrGluGluIleAsnGlyProPheIleGluGluLys	481	722	TCTATGTTTTAT-----AGAGCCCCAGTGCAGTGGAGATGGAATAAT	766	
Db	353	GATACGAAATAGTAGTATCATTTGGGAAGAAATAAACGGGCCCTTCATAGAGAGAG	412	675	AspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThr	694	
Qy	482	ThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSerPhe	501	767	GACAAGCATATGACCTGTGACTGGTCTCCAGGTGGGACCTACCACTTCGTTTT	823	
Db	413	ACTTACTTGCTTC-----TTTACTTTT	436	695	ValLysAspGlnGlnGlyLeuSerSerThrThrLeuThrValAlaVal	711	
Qy	502	ArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaLeuIleVal	521	824	-----TATTTCTACCTGGTGTTTTAAATCAATCTCTTTGCTTTTCAT	865	
Db	437	TTGTTTTGTTTT-----	448	712	-----LysLysGluAsnAsnSerProProArgAlaAArgAlaGlyGlyArgHis	727	
Qy	522	AsnAsnAlaValAspTyrProProValAlaAsnAlaGlyProAsnHisThrIleThrLeu	541	866	TTTGCTATATTGACTTAGAAAAATAATAGTCTCCAGAGCCCGGCTGGGACAGAT	925	
Db	448	-----	448	728	ValLeuValLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAsp	745	
Qy	542	ProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspHisGlnIleValLeu	561	926	GTTCTTGCTCTCCCAATATTCATTTGATGGTTCAAGTCTTAAGAGCCCAAG	985	
Db	449	-----GTTTTTATT	457	746	-----AspGlnArgIleValSerTyrLeuPheArgGlyGlnSerProAlaAla	763	
Qy	562	TyrGluTrpSerLeuGlyProGlySerGluGlyTyrHisValMetGlnGlyValGln	581	986	AATTGGGACCGAGGATTGACAGCTGCCCTGTGATCCGT-----GTTCTTTTCG	1033	
		---		764	GlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsnLeuValGlu	783	
		---		1034	CAGGATGTCATCGATGGCTCTGACACAGTGGCTCTGCAGCTTACAGATCTGGTGAG	1093	
		---		784	GlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAspThr	803	
		---		1094	GGGTGTACACTTTTCCACTTCTTATGCTCTAAAGCCTGTGCAAAATGTCGAAGTCAAAAT	1153	
		---		804	AlaThrValGluValGlnProAspProArgLysSerGlyLeuValGluLeuThrLeuGln	823	
		---		1154	CCTGTATCTTCTTTTCCCGACAGCCCTAGGAAGAGTGGCTGGTGGAGCTGACCTGCGAG	1213	
		---		824	ValGlyValGlyGlnLeuThrGlnArgLysAspThrLeuValArgGlnLeuAlaVal	843	
		---		1214	GTTGGTGTGGGAGCTGACAGAGCAGCGGAAGACACCTTTGTGAGGAGCTGGCTGTG	1273	
		---		844	LeuLeuAsnValLeuAspSerAspIleValGlnLysIleArgAlaHisSerAspLeu	863	
		---		1274	GGCGTGTGATGGTAGTCACCATCTGACCATGTG---GCTGCTCTCTCATTTGCTCTC	1330	
		---		864	---SerThrValIleValPheTyrValGlnSerArgProProPheLysValLeuLysAla	882	
		---		1331	CACAGACCGGATGTTGTTTATGTACAGAGCAGCGCCCTTTCAAGGTTCTCAAGCT	1390	
		---		883	AlaGluValAlaArgAsnLeuHisMet - ArgLeuSerLysGlyLysAlaAspPheLeu	902	
		---		1391	GCTGAAGTGGCCCGAAATCTGCATAGCTGAATCCAGGGATTCCTTTTCTGCTGCT	1450	
		---		902	uPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGlyHisGlyHi	922	
		---		1451	GCCTCAGCCCTACACACCTCTTT - GCAGGTGCTCTTGAAGTGTCTGAGCTGGTCA	1509	

Db 1962 GCAGGCCAGAGGACGAGCCAGCAGAGAAATCTCCAGAGGTCCATCTGAAGGT 2021
 Qy 339 aGlyAspAsnLeuIleThrLeuProAspAsnGluValGluLeuIleValAlaPheValAl 359
 Db 2022 GCGGAGTGAT-----GCGCTGCCGTAGTAAATGTGGCGAGAGCGCGTACAGT 2072
 Qy 359 aProAlaProValGluThrThrTyraSerThrGluThrAsnLeuIleSerHisProTh 379
 Db 2073 CGTCGCGAGCAGCGCTTGAACAGACCTATGCGCTGAGCGGAGCGGCGAGCCCGCGC 2132
 Qy 379 rAsp-----TyrgInGlyGluIleLysGlnGlyHisLysGlnThrLeuAsnLe 395
 Db 2133 GGACGAAGAGTGACCTATAAATGGCTGGTGGCAGCGGATCGTTCGCATCCCAATGC 2192
 Qy 395 uSerGlnLeuSerValGlyLeuTyValPheLysValThrValSerSerGluAsnAlaPh 415
 Db 2193 GGATCAGGCCCAAGCCCGCGGCGTGTGCCGAG-----AATGCGAC 2234
 Qy 415 eGlyGluGlyPheValAsnValThrVal-----LysProAlaArgArgVa 430
 Db 2235 GGGTGAGGGCGCTCATGAGCTGACCGCTCACGAACGAGTCGGGCAAGACGCGGAGGCGAC 2294
 Qy 430 lAsnLeuProValAlaValSerProGlnLeuGlnGluLeuThrLeuProLeuTh 450
 Db 2295 G-----GTCMAGTACCGCGGTGGCGCCGAGTC----- 2325
 Qy 450 rSerAlaLeuIleAspGlySerGlnSerThrAspAspThrGluIleValSer----- 467
 Db 2326 ---ACGATCAGCGCGGTACGCGAGCAGGTGAAGGACAAACGAGCAGCGAGCTACAGCGC 2381
 Qy 468 -----TyrHisTrp----- 471
 Db 2382 GAATGCGAACTTCGACGCGCGGACCTACTCTGGAAGCTGTCCAGCGTGTCCGCGCAAGCA 2441
 Qy 471 uGluIleAsnGlyProPheIleGluGlyLeuThrSerValAspSerProValLeuArgLe 491
 Db 2442 GCGCGGACGCGGTAGCGGGCGCAATGTGAACACCGCGTG----- 2481
 Qy 491 uSerAsnLeuAspProGlyAsnTySerPheArgLeuThrValThrAspSerAspGly-- 510
 Db 2482 ---GGTCTGATGCGGCGACGTACACGCTGCTGTCGTCGACGACGTAACGCGCGCGCGA 2537
 Qy 511 -AlaThrAsnSerThrThrAlaAlaLeuIleValAsnAsnAlaValAspTyPro---Pr 529
 Db 2538 CCGGAGCGGCGCGGACGAGAACTGACCGTTGAGCGCACGCTAGCGCTTCCGCGACC 2597
 Qy 529 oValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsn-----SerIl 546
 Db 2598 GGAGGCCAACCGCGCGCGGACCGCGTGTGGCGACGACACCGCGGTGTGGCTA 2657
 Qy 546 eThrLeuAsnGlyAsnGlnSerAspAspHis---GlnIleValLeuTyThrTrpSe 565
 Db 2658 TCCGCTGTCGCGAGCGCAAGTCGCGACCCGCAACCGGCGAGCGCGATCATCAATGGCG 2717
 Qy 565 rLeu-----GlyProGlySerGluGlyLysHisValVa 576
 Db 2718 CGTGTGTGCGGCGCGTTTCACGTTGCGCAACGCGTGTGTCGCGGATGCGGAGCGGATCGT 2777
 Qy 576 lMetGlnGlyValGlnThrProTyThrLeuHisSerAlaMetGlnGluGlyAspTyThr 596
 Db 2778 ACCGAAGGCGAC----- 2804
 Qy 596 rPheGlnLeuLysValThrAspSerArgGlnGlnSerThrAla***-----ValTh 614
 Db 2805 GTTCGAGCTGAAGTGAGCAACCGCGCGGAGACGCGCGGAGCGGACGACGAGGTGAC 2864
 Qy 614 rValIleValGlnProGluAsnAsnArgProProValAlaValAlaGlyProAspLysGl 634
 Db 2865 CGCGATCGCG-----CCGCGCGTGACGCTCAGCGGC----- 2895
 Qy 634 uLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSerSerAsp----- 651

Db 2896 -----AGCAGCACCGTGCACCGGCGAGCGGGCAGCCAGCTACAGGC 2936
 Qy 652 -----AspHisGlyIleValPheTyHisTrp-----GluHisValArgGlyProSe 667
 Db 2937 GAACCGCAACTTCGAGGAAGCGACCTATAGCTGCGACTCACCAGTACGTCGGCGCAGCA 2996
 Qy 667 rAlaValGluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnValGl 687
 Db 2997 GGTGTATCGGCGAGCGGTGCGGATGGAACCTGTCC-----GCTACGTGAAGAAGCG 3050
 Qy 687 yThrTyHisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrSerThLe 707
 Db 3051 CAATACACCTCAAGTGTCTCGCATCGCTCAAGCGCGAGCGTAGC---GCGACGCG 3107
 Qy 707 uThrValAlaValLysLysGluAsnAsnSerProProArg-----AlaArgAlaGl 724
 Db 3108 GACCTCGCGCTGAAGTGGAGCGCGCGTGGCGCCCAAGCCGCTGGAGGCCAAATCAG 3167
 Qy 724 yGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAspGlySerArgSerTh 744
 Db 3168 CGGT---CCGTTGAATATCAGGCGGCGCAGCATCTTACTTGTGACCGTCCGAGCGCAATCGT 3224
 Qy 744 rAspAspGlnArgIle---ValSerTyLeuTrpIleArgAspGlyGlnSerProAlaAl 763
 Db 3225 CAACCCGAAACGAGTACGCGACCTACTTGTGACCGTCCGCGCGCAG----- 3273
 Qy 763 aGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThr-----As 780
 Db 3274 -----CTGGCTCGAATGCGGAAACCGCGAGCGTCCGTTGACGCGCGCTGCCAA 3323
 Qy 780 nLeuValGluGlyVal---TyrThrPheHisLeuArgValThrAspSerGlnGly---Al 798
 Db 3324 TGCCACAGTGGCGCGCGAGTACCATCAAGCTGTGCTGTACGACGACGAGAAACAGCT 3383
 Qy 798 aSerAspThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuVa 818
 Db 3384 GTCCAGTACGCGAGCTACACGCTGACGCTGAAGCGCGGAAACACCGCGCGCGCCAGC 3443
 Qy 818 lGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuVa 838
 Db 3444 CTATCCGCGCTACAAGCGCGT-----ACGAGTACAAGCGTGTGAGCTGCT 3491
 Qy 838 lArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLysIleAr 858
 Db 3492 GCT-CAACTACGCTACGCTCTATCAGTG-----CAAGCGGTCCCGTACCGCGCTGCT 3544
 Qy 858 gAlaHisSerAspLeuSerThrValIleValPheTyThrValGlnSerArgProPheLy 878
 Db 3545 -GCA-----GTACGCGCGCGCGCGCTATGAG 3570
 Qy 878 sValLeuLysAla 882
 Db 3571 CCGGTAAGGCT 3583
 RESULT 15
 US-10-278-698-1034
 ; Sequence 1034, Application US/10278698
 ; Publication No. US20050037344A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PathoArray GmbH
 ; APPLICANT: Stuhlmueller, Bruno
 ; APPLICANT: Haupt, Thomas
 ; TITLE OF INVENTION: Nucleic Acid Array
 ; FILE REFERENCE: 030027US
 ; CURRENT APPLICATION NUMBER: US/10/278,698
 ; CURRENT FILING DATE: 2002-10-23
 ; NUMBER OF SEQ ID NOS: 1050
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 1034
 ; LENGTH: 82027
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-278-698-1034

Alignment Scores:			
Pred. No.:	1.74e-07	Length:	82027
Score:	225.00	Matches:	252
Percent Similarity:	33.64%	Conservative:	186
Best Local Similarity:	19.35%	Mismatches:	416
Query Match:	4.03%	Indels:	448
DB:	21	Gaps:	69
US-10-764-390-3 (1-1072) x US-10-278-698-1034 (1-82027)			
QY	78	TyrLeuValSerCysProHisIysGluAsnCys---	GluProLysLysMetGlyProIle 96
DB	41174	TTCCAGGGTCAGTCCGCAAAATAATATGTTGTTAGGGAAGGCGCTCAAAATCGGAGCCAATT	41233
QY	97	ArgSerTyrLeuThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyr	116
DB	41234	GTTCGGAGACATCCATTGATGCTGATGCTGCCCCACCTCCCAATATTGTGGAT---	41290
QY	117	GlyAspMetLeuAsnArgGlySerProSerGlyIleTyrGlyAspSerProGluAsp	136
DB	41291	GTGACACAGATTGAGTATCTTAACCTTGACACTGACCCC	41329
QY	137	IleArgLysAspLeu***PheLeuGlyLysAspTyrGlyLeuGluGluMetSerGluTyr	156
DB	41330	-----AAGAAAACCTGGTGTCTTCCAAATTACAGGGTAT	41362
QY	157	***AspAspTyrArgGluLeuGluLysAspLeuGlnProSerGlyLysGlnGluPro	176
DB	41363	CATCTCGAGTTCAAGGAA---AGAAACAGCCCTTTTGTGGAAGAGAGCTAAACAAGACTCGG	41419
QY	177	ArgGlySerAlaGluTyrThrAspTyrGlyLeuLeuProGlySerGluGlyAlaPheAsn	196
DB	41420	ATAAGATCAGAGACTTTAAAGTCACAGATTAACTGAAGGTTCTTGAATATGAATTCGA	41479
QY	197	-----SerSerValGlyAspSerProAlaValProAlaGlu-----	208
DB	41480	GTTATGGCAATCAATTTAGCAGGTGTGGGC---AAGCCAAAGCTACCATCAGAGCCTGTT	41536
QY	209	ThrGlnGlnAspProGluLeuHisTyrLeuAsnGluSerAlaSerThrProAlaProLys	228
DB	41537	GTGGCCTCGACCCCA-----ATTGATCCTCTCGGA	41566
QY	229	LeuProGlu-----ArgSerValLeuLeuProLeuProThrProSer	243
DB	41567	AAACCTGAGTTATTAAACATAACAGGAATTCAGTGACTCTCATTTGGACTGAACCTAAA	41626
QY	244	-----SerGlyGluValLeuGluLysGlu-----	251
DB	41627	TATGACGGTGGTCATAAGTTAACTGGATATATAGTGGAGAAGCAGATCTACTCTCGAAG	41686
QY	252	-----LysAlaSer-----	254
DB	41687	TCTTGGATGAAAGCCCAACCATGTTAATGTCCAGAAATGTCCTTTACTGTAACGACCTT	41746
QY	255	-----GlnLeuGlnGluGlnSerAsnSerSerGlyLysGluVal	268
DB	41747	GTTGAGGGTGGAATAATATGAATTTCAGAAATAGAGCAAGAAATACAGCAGGT---GCTATC	41803
QY	269	LeuMetProSerHisSerLeuProPro-----AlaSerLeuGluLeuSerSerValThr	286
DB	41804	AGTGCTCCATCAGAAAGTACAGAAACCATTTATTGCAAGGATGAATACGAGGCACCAACA	41863
QY	287	ValGluLysSerProVal-----LeuThrValThrProGlySerThr-----	300
DB	41864	ATTGCTCTTGATCCCAATAAAGATGGGCTAACAAATTAAGAGCGGGATACCAATTGTT	41923
QY	301	-----GluHisSerIleProThrProProThrSerAlaPro	313
DB	41924	TTGAATGCCATTAGCAATTTCTGGCAAAACCCCTTCCAAATCAAGTTGGTCCAGGCAGGA	41983
QY	314	SerGluSerThrProSer-----GluLeuProIleSerProThrThrAlaProArg	330

Db	41984	AAAGACATTAGACCATCAGATATCACTCAGATAACTTTCAACCCCAACATCTTCCATGCTT	42043
QY	331	ThrValLys-----GluLeuThrValSerAlaGly-----	340
DB	42044	ACTATCAAGTATGCCACTAGAAAGATCGGGTGAATATACCATCAGTCTACCATCTCT	42103
QY	341	-----AspAsnLeuIleIleThrLeu-----ProAspAsn	350
DB	42104	TTTGGCAGGAAGTGGAAACATGTGAAGTAAACAGTCCCTTGATGTACCTGCTCCCGCAGGT	42163
QY	351	GluValGlnLeu-----LysAlaPheValAlaProAlaProProVal	364
DB	42164	CCTGTTGAAATCAGTAATGTTTCTGCTGAAAGCAACACTTACATGTGACACCTCCCTTG	42223
QY	365	Glu-----ThrThrTyrAsnTyr	370
DB	42224	GAAGATGGCGCTCACCAATTAAGTCTATATATCTTGAAGAGAGAAACAGCGGACTT	42283
QY	371	GluTrpAsnLeuIleSerHisProThrAspTyrGlnGly-----Glu	384
DB	42284	TTGTGGACAGTGGTTTCTGAA-----GATATTCACTCTTCAGGCATGTGGCAACCAA	42337
QY	385	IleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrVal	404
DB	42338	CTTATCCAAGGAATGAG-----TACATC	42361
QY	405	PheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrVal	424
DB	42362	TTCCGG-----GTCTCAGCTGTAAACACATATGGCAAGGA---GAACCTGTACAGTCT	42412
QY	425	LysProAlaArgValAsnLeu-----ProProValAlaValValSerProGlnLeu	442
DB	42413	GAACCTGTCAAAATGTGTAGACAGATTTGGTCCCTCTGCGCCCTCTGAAAAACAGAGGTA	42472
QY	443	GlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThrAspAsp	462
DB	42473	TCAAAATGTCACT-----AAGAACACTGCCACTGTGAGTGGAAAGCCAGTGGATGAT	42526
QY	463	-----ThrGluIleValSerTyrHisTyrGluGluIleAsnGlyProPheIleGluGlu	480
DB	42527	GGTGGCAGCGAAATTACAGGATATCATGTAGAAAGGAGAGAAAGAAAGCCTCGCATGG	42586
QY	481	LysThrSerValAspSerProVal-----LeuArgLeuSerAsnLeuAspPro	496
DB	42587	GTGAGACCAATAAAACACACAGTTTCCGATCTCAGGTGCAAAAGTAACAGGACTGCAAGA	42646
QY	497	GlyAsn---TyrSerPheArgLeuThrValThrAsp-----SerAspGlyAlaThrAsn	513
DB	42647	GGAAGACCTACGAATTCGGTGTGTCAGTGCAGAAACACAGAGGAAATTTGGTCCACCCAGT	42706
QY	514	SerThrThrAlaAlaLeuIleValAsnAsnAlaValAspTyrProProValAlaAsnAla	533
DB	42707	GAGGCTTCAGATTCTGTTCTGATGAAGATGCAGCA---TATCTCTCAGGACCCTTCA	42763
QY	534	GlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSer	553
DB	42764	AATCGCATGTCACTGATACCTACCAAGAAATCTGCTTCTTGGCATGGGCAAGCCTCAT	42823
QY	554	SerAsp-----AspHisGlnIleValLeuTyrGlu	563
DB	42824	TATGATGGTGGACTTGAATCACTGGCTATGTCGTGGAGCATCAAAAAAGTAGAGACGAG	42883
QY	564	---TTP-----	564
DB	42884	GCCTGGATAAAGATACACAGGAACCCCTCTCAGAAATCACTCAGTTCTGTTCTGCTGAT	42943
QY	565	-----SerLeuGlyProGlySerGlyGly	572
DB	42944	CTTCAGACTAAAGAAAATACAACTTCAGAATCAGTGCCATCAACAGATGCAGGTGTGG	43003
QY	573	LysHisValValMetGlnGlyValGln-----ThrProTyrLeuHis	586
DB	43004	GAGCCAGCGGTGATTTCCAGATGTTTGAATCTGTAGAACGGGAGATGGCTCTCTGATTTGAA	43063

```
Qy 587 LeuSerAla-----MetGlnGluGly----- 593
Db 43064 CTAGATCCGAGCTTCGAAGAACAACCTTGTGTGTAGACGAGCACTCAGTATTAGGATATTT 43123
Qy 594 -----AspTyrThrPhe-----GlnLeuLysVal 601
Db 43124 GTGCCAATTAAAGTCGTCCTGCTGCTGAAGTGCATGGACCAAGATAACATCAACCTG 43183
Qy 602 ThrAspSerSerArgGlnSerThrAla***ValThrValIleValGlnProGluAsn 621
Db 43184 AAAAACCGAGCAACATTGAAATAACGGAATCAATTTACTCTCTCTGATTATCCAGAAATGT 43243
Qy 622 AsnArg-----ProProValAlaValAla 629
Db 43244 AACAGATATGATACCGGTAAATTTGTTCATGACCAATTGAAACCCCGCTGGGAAGAAAGT 43303
Qy 630 GlyProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSer 649
Db 43304 GGCCTTTGAACGTCAGAGTCTTGCACACGCCAGCCAGTCTCTCAACCTCGCGCTTACA 43363
Qy 650 SerAspAspHisGlyIleValPheTyrHisTrpGlu-----HisValArgGlyProSer 667
Db 43364 GACATCAAAAGGACAGTGCACCTGCGACCTCCCTCTCTGATAGATGGAGGTCA 43423
Qy 668 AlaValGluMetGluAsnIleAspLysAlaIleAlaThr----- 680
Db 43424 CGTATAACAACTACATTGTAGAGAAACGCTGAAGCAACACCGGAAATCTTATTCACAGCC 43483
Qy 681 -----ValThrGlyLeuGlnValGly-----ThrTyr 689
Db 43484 ACCACTAAGTCCATAATGCACATATATAAGTTTACCGCTTGTCTGAAGGTGTGAATAT 43543
Qy 690 HisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThrVal 709
Db 43544 TTCTTCAGAGTGATGGCAGAGATGAATATGGAATTTGTGTAGCCCAACAGAACTACAGAG 43603
Qy 710 AlaValLysLysGluAsnAsnSerProProArgAlaArgAlaGlyArgHisValLeu 729
Db 43604 CCGGTAAAGCCTCTGAAGCACCATCTCCA----- 43633
Qy 730 ValLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIle 749
Db 43634 -----CCAGACACCTTAAATCATGACATAACTAAGACACC----- 43672
Qy 750 ValSerTyrLeuTrpIleArg-----AspGlyGlnSerProAlaAlaGlyAspVal 766
Db 43673 GTACGCTGGCATGGCCTTAAGCCCAACACCATGTTGTGCAGCAAGATCACTGGCTATGTG 43732
Qy 767 IleAsp-----GlySerAsp-----HisSerValAlaLeuGln----- 777
Db 43733 ATTGAAGCCCAAGAAAGGCTCTGACCACTGGACCCACATCACAAACCGTGAAGGGTTA 43792
Qy 778 -----LeuThrAsnLeuValGluGly---ValTyrThrPheHisLeuArgValThr 793
Db 43793 GAATGTGTGTGGAGAACTTAACCTAGAGGAGGAATATACCTTCCAAAGTATGGCGATG 43852
Qy 794 AspSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnPro-----Asp 811
Db 43853 AACAGCGCGGGAGAGT-----GCCCTAGAGAAAGCAGACCCGCTCATTTGTC 43900
Qy 812 ProArgLysSerGlyLeuValGluLeuThrLeuValGlyValGlyGln----- 828
Db 43901 AAGGAGCAGACAATGCTTCCAGAGCTGGATCTCGT---GGCATCTATCAGAAATCGGTC 43957
Qy 828 ----- 828
Db 43958 ATTGCCAAGCTGGTGCAACAACATCAAGTTGAAATTCAGTCTCGGTGCGACCGAAGCCC 44017
Qy 829 ---LeuThrGluGlnArgLysAspThrLeuValArgGln----- 840
Db 44018 ACAGTGACATCGAANAAGGAGACCAAAATTTCTTAACACAGACAGAGAGTTAATTTGAA 44077
```

```
Qy 841 -----LeuAlaValLeuLeuAsnVal-----LeuAspSerAspIleLysVal 854
Db 44078 ACCACAGCACTTCAACCACTTTAAATATCAATGAGTGTGCAGAAAGTGAATGGGCCCC 44137
Qy 855 GlnLysIleArgAlaHisSer-----AspLeuSerThrValIleValPheTyrVal 871
Db 44138 TATCCATTAAACAGCAAGGAACATTGTAGGAGAGGTGTGTGATGTCATCACCATCAAGTC 44197
Qy 872 GlnSer-----ArgProPheLysValLeuLysAlaAlaGluValAlaArgAsnLeu 889
Db 44198 CATGATATCCAGGCCACCTACTTGGACCAATCAAAATTTGTATGAAGTTTCA----- 44248
Qy 890 HisMetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAsp 909
Db 44249 -----TCTGATTTTGTAACTTCTCT----- 44269
Qy 910 ThrAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArg 929
Db 44269 ----- 44269
Qy 930 CysIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyGlu 949
Db 44270 -----TGGGACCCCACT 44281
Qy 950 SerAsn-----CysGluTrpSerIlePheTyrValThrValLeuAlaPheThrLeu 966
Db 44282 GAGAACGATGTTGTGTACCAATGAACAACTATGATGTGGAATGCGGCAG----- 44332
Qy 967 IleValLeuThrGlyGlyPheThrTrpLeuCysIleCysCysLysArgGlnLysArg 986
Db 44333 -----ACTGACAGTACTACCTGGTGTGAGTTAGCA-----ACC 44365
Qy 987 ThrLysIleArgLysLysThrLysTyrThr---IleLeuAspAsnMetAspGluGlnGlu 1005
Db 44366 ACCGTTATACGTACTACTATAAAGCCACCCGCTTACTACTGTGATTAGATCAGTTC 44425
Qy 1006 ArgMetGluLeuArgProLysTyrGlyIleLysHisArgSerThrGluHisAsnSerSer 1025
Db 44426 CGGTAAAGCTCAGATAGATATGGAGTTGGACCGAGCATCACA-----TCAGCA 44476
Qy 1026 LeuMetValSerGluSerGluPhe----- 1033
Db 44477 TGCATAGTTGCCAACTATCCATTTAAGGTTCTCGACCTCTGTTACCCCTCAGGTAAC 44536
Qy 1034 -----AspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnPro 1051
Db 44537 GCAGTTACCAAGGATTCAAATGACAATTAGCTGGCATGAGCCACTTTCTGTGTTGGAAGC 44596
Qy 1052 LysVal-----SerMetAsnGlySerIleArgAsnGlyAlaSerPheSerTyrCys 1068
Db 44597 CCCATTTTAGGATATCATGTTTGAAGAAAGAACGAAATGTTTCTCTGGCAGACTGTG 44656
Qy 1069 SerLys 1070
Db 44657 AGCAAA 44662
```

Search completed: October 12, 2005, 16:49:57
Job time : 4109 secs

THIS PAGE BLANK (USPTO)

513	Qy	AnSerThrAlaLeuLeuValAsnAsnAlaValAspTyrProProValAlaAsn	532
		: : : : : : :	
431	Db	AAATCAACCACCTGCCACGATAGCAGCTTTAAAGGAGACCGATTATGTCGGGTGGCAAT	372
		: : : : : : :	
533	Qy	AlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGln	552
		: : : : : : :	
371	Db	GCGGGGATGCCGTGATCTTGTTTGGCCAAACAACTAATGTCACTCTGAATGCGACAGCC	312
		: : : : : : :	
553	Qy	SerSerAspAspHisGlnIleValLeuTyrSerLeuGlyProGlySerGluGly	572
		: : : : : : :	
311	Db	AGTTGGATGATCACGAGATCGTTCTTGGAGTGGACCAAGGATGCTAGCGACGAGGCC	252
		: : : : : : :	
573	Qy	LysHisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGlu	592
		: : : : : : :	
251	Db	AAGGGTGTGGATTCGAGAATCAAGAACACCCCTATGTTCAAGCTGTCTCAATTTTGGAGGAG	192
		: : : : : : :	
593	Qy	GlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla**	612
		: : : : : : :	
191	Db	GGCATGTACACTTTTGTCTAAAGTAACCGATGGCAGTGGCAATCGAGTACAGCTAAG	132
		: : : : : : :	
613	Qy	ValThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaGlyProAsp	632
		: : : : : : :	
131	Db	GTCCATGTGTTTGTGAAGCCTCCCAAGAAATCTCCACCAGTTGCTGAGGCGAAGGTAAT	72
		: : : : : : :	
633	Qy	LysGluLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSerAspAsp	652
		: : : : : : :	
71	Db	ACGACTACTAGCTTGCCCAATTAAATGGGGTCTTTTGAATGGCTCCGATTGCAAGGACGAC	12
		: : : : : : :	
653	Qy	HisGlyIle	655
11	Db	ATTGGCATC	3

RESULT 2

```

US-09-270-767-16109/c
; Sequence 16109, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16109
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16109

```

Db 1014 TGTGG-GCTTCC-----ATGCTCTGG-----GAGTGTAGTGTCTGCTC 1051
Qy 79 LeuValSerCysProHisLysGluAsnCysGluProLysLysMetGlyProIleArgSer 98
Db 1052 CTCAGGCCACCCCTTCCAGGAGCCAGGACCGCATGAGCACTACCTGGTGGCCGGG 1111
Qy 99 TyrLeuThrPheValLeuArgPro----- 106
Db 1112 TACAAGCCAGCCCGCTCTCCCCACCTTCTCGAGGCGCGAGCGGTGCTCGCGGTC 1171
Qy 107 -----ValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetLeuAsnArgGly 124
Db 1172 ACCGCCGCGAGCCCGCGCGCATACAGCGCTTCTCCAGGCCCTTCCCAAGCGGGT 1231
Qy 125 SerProSerGlyIleTrpGly-----AspSerProGluAspIleArgLysAspLeu 141
Db 1232 GCCGCATGGCGGAGGCGCGGACCGTGACTCCATGACCAACCGTGGCGGTGGAG 1291
Qy 142 ***PheLeuGlyLysAspTrpGlyLeuGluMetSerGluTyr***AspAspTyrArg 161
Db 1292 AGCTTCGCGCGGCGCACTTC---ATCAGGTCACTCCCTGCGGCTCGATGACGCCCTG 1348
Qy 162 Glu---LeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAla 180
Db 1349 GAGGCTCGGCACAGGAC-----GGCGTGCTC 1375
Qy 181 GluTyrThrAspTrpGlyLeuLeu-----ProGlySerGlu 192
Db 1376 GACGTGGCGAGACGGCGCTGCTGGGTGAGCGTGGCGCAACGTGGCGCGGTTCGCTG 1435
Qy 193 GlyAlaPheAsn----- 196
Db 1436 GCGCCCTTCTCCGCAACCGTCTCCACCACAGCAGCAGCGCCACGCTGGAGTTCCCTCG 1495
Qy 197 -----SerSerValGly----- 200
Db 1496 GGCACACCCCTGACCTTCAGTCTGATGGCGCGCGCGGCGGCGGCGCCACCGCTCCCG 1555
Qy 201 -----AspSer 202
Db 1556 GTGAGGTGGCGCGCGCGAGCCGACGACACACGCGCGCTGCGCATCCACTTCGACGAG 1615
Qy 203 ProAlaValProAla-----GluThrGlnGlnAspProGluLeuHisTyr----- 217
Db 1616 CCCTCGTCTCCCGCGCAGCGGAGCAGCAGCGTTTGACCTCGGGTGAATACGACGAA 1675
Qy 218 LeuAsnGluSerAlaSer-ThrProAlaProLysLeuProGluArgSerValLeuLeuPr 237
Db 1676 GTGCCGCGCGCGCGGCGGAGGACGACTTCGAGGCGCGCTGTCTCATGGACCTCGTCC 1735
Qy 237 oLeuProThrThrProSer-----SerGlyG1 246
Db 1736 AGTGGCTCAACCCCTACGCGGAGCGGAGCTGGAAGTTCAGGGTCCGAGGGCGGGCGC 1795
Qy 246 uValLeuGlyLysGluLysAlaSerGlnLeuGlnGlnSerSerAsnSerSerGlyLys 266
Db 1796 TACATGACGCGCGCCCAACCGCGCATGACAGCGGACCTTCTCTCATCTCCCTGGATG 1855
Qy 266 sGluValLeuMetProSerHisSerLeuProProAlaSerLeuGluLeuSerSerValTh 286
Db 1856 AAGTGAATGAGCCGGG-----CCTTCTCGTTCAGCTTCCGTACCGC 1900
Qy 286 rValGluLysSerProValLeuThrValThrProGlySerThrGluHisSerIleProTh 306
Db 1901 CACTCGTTCGAGTCCG-----ACATCTACCCCGCGGCT-----CCATC 1939
Qy 306 rProProThrSerAlaAlaProSerGluSerThrProSerGluLeuProIleSerProTh 326
Db 1940 TTCCCGTACTACGAGCGCGCGTCTCGAGTTCACCCACG---ACGCGCTGCAGTGGTAC 1996
Qy 326 rThrAlaPro-----ArgThrValLysG1 334
Db 1997 GACATCGCGCGATGGGCGTGGCTCTCATCTCTACCTCGACCCGACGAGCGGCTCCCGC 2056

Qy 334 uLeuThrValSerAlaGlyAspAsnLeuIleIleThrLeuProAspAsnGluValGluLe 354
Db 2057 GGCTACATCAGCAGGGGC----- 2075
Qy 354 uLysAlaPheValAlaProAlaPro-ProValGluThrThrTyrAsnTyrGluTrpAsnL 374
Db 2076 -----CCACCCCTGATGTACCGCAACGC----- 2100
Qy 374 eulleSerHisProThrAspTyrGlnGlyGluIleLysGlnGly----- 388
Db 2101 --GTGGGTGGCCAGAACGTCACGACGAGAACCGCAAGACGGTGGCTTCCCGAGTGG 2158
Qy 389 --HisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrVal---PheLysV 407
Db 2159 TCCACCGCCGAG--CTGCGGCTGGCGGACATGTTCGCCGCGCAAGAACGTCCAGTTCGCT 2216
Qy 407 alThrValSerSerGluAsnAlaPheGlyGlu-----GlyP 419
Db 2217 TCGCATTTGTACCGACACCGCGGTGGCGGTACGGCTTCGACCTGGACGACGTGGCT 2276
Qy 419 heValAsnValThrValLysPro---AlaArgArgVal----- 430
Db 2277 TCACCAAGTGGAGCGCATGCTTTCAGCGCGCGGTGGAGCAGCAGCGCGCGCA 2336
Qy 431 -----AsnLeuProProValAlaVal-----SerProG 441
Db 2337 CCTGCAACCAACCGCGCGTGGCCCAACCGCGCGCGACGTGGCAGCGCCCGCGG 2396
Qy 441 lnLeuGlnGluLeuLeuLeuProLeuThrSerAlaLeu-----lleA 455
Db 2397 TGTATGACGGGTGAGGAACCCCGAGACGGGGGACCTGGAGCTCGTGGCCACACGTGG 2456
Qy 455 spGlySerGlnSerThrAsp---AspThrGluIleValSerTyrHisTrpGluLeuA 474
Db 2457 ACGGCACCGCAGCTATGATCCCGGAGGCGCAGCCCTCACCTACCTACCTGGACGAGTGA 2516
Qy 474 snGlyProPheIleGluGluLysThrSerValAspSerProValLeuArgLeuSerAsnL 494
Db 2517 GCGGG-----CGGCGGTCACTCGTGTGAACG 2543
Qy 494 euAspProGlyAsnTyrSerPheArgLeuThrValThrAspSer----- 508
Db 2544 CGGACACGGCCACCCCGTCTGTTCAGCGGGATGTCGCGACGACTCATCTCAGGTCC 2603
Qy 509 -----AspGlyAlaThrAsnSerThr-----ThrAlaAlaLeuIleV 521
Db 2604 AGCTCGTGGTGGCGATGGCTGGAGACCGCGGCGCACACAGGTCCGGGCTGGGTCC 2663
Qy 521 alAsnAsnAlaValAspTyrProProValAlaAsnAlaGlyProAsnHisThrIleThrL 541
Db 2664 TCAAC-----GTCAACCGCGCCCGCCACCGCGGTGGTGAACG 2699
Qy 541 euPro-----GlnAsnSerIleThrLeuAsnGlyAsnGlnInserS 554
Db 2700 CGCGCGCGAGGTGGCGGAGTTCCTCGGAGCCCGGTGACGCTGGACCCCGCGCTCGA 2759
Qy 554 erAsp---AspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyL 573
Db 2760 GCGACCGCGAGCGGGAGCTCATCTCTACAAGTGGCG-----CAGACGTCGGTCC 2810
Qy 573 yHisValValMetGln-----GlyValGlnThrProTyrLeuHisLeuSerAlaMetG 591
Db 2811 CCGCGTGGCGCTCAAGCGGGCTTCAGCGCGCATCGCGGAGTTCGAAGTGGCGGAGTGG 2870
Qy 591 ln---GluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnInserT 610
Db 2871 CCGTCGATCTACCTTACCTTTCAGTGTGGTGAACCTACGCGGACCGAGCAGAGA 2930
Qy 610 hrAla***ValThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaG 630
Db 2931 CGGCCACCGTGACGTGATGTGGCG---AACGTTGACCGCGCGCGCTGGCGATGCGG 2987

QY 630 lyProAspGlyGluLeuIlePheProValGluSerAlaThrLeuAspGlySerSers 650
 Db 2988 GC---TCCGAGCAGGTGGTGGACCGCTGGCCACCGCGCTCAGCGCGGTGGTG 3044
 QY 650 erAspAspHisGlyIleValPheTyrHis-TrrGluHisValArgGlyProSerAlaVal 669
 Db 3045 GACCGGACGGGATGCCCTACCTGTACGTGGCTGCGTGGAGGGCCCC---CCGGTG 3101
 QY 670 GluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnVal----- 686
 Db 3102 GAGTGCAGGGAGCGCCACGACCGCCACCGCTGTTCCACCGCCCCCGGGCGGAG 3161
 QY 687 GlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThr-----Ser 705
 Db 3162 ACGGTGCTGGCTTCGAG---GCCATCGCCAGCGCGAAGCGGTCTCCAGCGCCCCGATG 3218
 QY 706 ThrLeuThrValAlaValLysLysGluAenSerPro-----ProArg----- 720
 Db 3219 ACGCGCAGGTGACGGTGGCGCAGCCCAACAGGGCGCCAGGCCCTTCCGGCGTGGTG 3278
 QY 721 -----AlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeu 738
 Db 3279 GGTGACGCTCGGTGGGGCGCGAG-----GTGACGCTC 3311
 QY 739 AspGlySerArgSerThrAsp---AspGlnArgIleValSerTyrLeuTriPleArgAsp 757
 Db 3312 GACGGTTCGGCTCCAGCGACCGCGATGGCGAGCGCCCTCACCTTCGGGTGGAGCAGGTG 3371
 QY 758 GlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGln 777
 Db 3372 GCGGACCGTCCGCTCACCTCTCCGCTCCGAGGGTGCG-----CAGGTGACGTTCCGC 3425
 QY 778 LeuThrAsnLeuValGluGlyVal---TyrThrPheHisLeuArgValThrAspSerGln 796
 Db 3426 GCGCCGAGGTGAAGCAGCGCGTGAAGTGGTGCATTCGAGCTCACCGCGACGCGCGCGC 3485
 QY 797 GlyAlaSerAspThrAspThrAlaThrValGluValGlnPro 810
 Db 3486 GCGCGCGCCAGCAGCGCGTGTCTCGGGTGACGGTGCGCGG 3527

RESULT 4

US-09-902-540-2590
 ; Sequence 2590, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 2590
 ; LENGTH: 4287
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-2590

Alignment Scores:
 Pred. No.: 1,38e-18 Length: 4287
 Score: 317.00 Matches: 162
 Percent Similarity: 38.50% Conservative: 84
 Best Local Similarity: 25.35% Mismatches: 232
 Query Match: 5.68% Indels: 161
 DB: 4 Gaps: 28

US-10-764-390-3 (1-1072) x US-09-902-540-2590 (1-4287)

QY 184 AspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGlyAspSerPro 203

Db 2530 GACCGTCCATCGCGCGGGTGAACCTGGTGGCTCTACGGTTCCGCTGCCGATCCGGAA 2589
 QY 204 AlaValPro-----AlaGluThrGlnGlnAspProGluLeuHisTyrIleu 218
 Db 2590 GCGCAGCGGGTGACCTTTCACCTGGGCAAGAG---TCTGGCCCTGGCGTCACTCGCG 2646
 QY 219 AsnGluSerAlaSerThrPro-----AlaProLysLeuProGluArgSerValLeu 235
 Db 2647 GGTCCCAACACCTGAACCCGCTGCTTACGGCGCGCAGGTTACCGAGTCCACCGCATG 2706
 QY 235 ----- 235
 Db 2707 GTGTTTCAGCTGACCGTCTCCGACGGGTCAAGAGCTCCACGGAAGTCCGCTGACCGTGACG 2766
 QY 236 LeuProLeuProThrThrProSerSer-----GlyGluValLeuGluLys 250
 Db 2767 GTGGCGCTCCCAACAACTCCGCCACCGTGAACCGGGCCTCGACGGCATCGTCGAGGAG 2826
 QY 251 GluLysAlaSerGlnLeuGlnGlnSerSerAsnSer----- 263
 Db 2827 CGCGTGAGTACACGCTGAGCGGCTCCGCCAGCATGGCGATGGCAATGCCCTCACGTAC 2886
 QY 264 -----SerGlyLysGluValLeuMetProSerHisSerLeuProProLa 278
 Db 2887 CTGTGGACCCAGGTCTCCGGTACCGCGTGGCGGTGAAGGACTACACCAACCGCGCGCG 2946
 QY 279 SerLeuGluLeuSerSerValThrValGluLysSerProVal-----LeuThrValThr 296
 Db 2947 ACGTTCATCGCGCGGAGGTGACGTCGTGATGAGTGGTCTTCCGCTGACGGTCAAGC 3006
 QY 297 ProGly-Ser-----ThrGluHisSerIleProThrProProThrSerAlaAl 312
 Db 3007 GACGCGATCGCCACCGGTGAACACACGCTCACCCTGACGGTCAACCAACCGCGACCGCG 3066
 QY 312 aProSerGluSerThr-----ProSerGluLeuProIleSerProThrThrAlaAl 329
 Db 3067 CCCATCGTCTCCGACACGCTCGCGTTCGCCCGCGGCACGTGTCCACCGTACGCGCTCC 3126
 QY 329 aArgThrValLysGluLeuThr-----ValSerAlaGlyAspAsnLeuIleIle--Thr 346
 Db 3127 GCGGTTCGACCGCGATGGTGACGCTCAGCTACAGCTGGGAGCAGACGGGTGGTTCGACG 3186
 QY 347 LeuProAspAsnGluValGluLeuLysAla-----PheValAlaProAlaProVal 364
 Db 3187 GTCCGCTCAATGGCGCGCAGCAGTCCGCCATCTCTTCCGCCACCCCGTCCCG----- 3240
 QY 365 GluThrThrTyrAsnTyrGluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGlu 384
 Db 3240 ----- 3240
 QY 385 IleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrVal 404
 Db 3241 -----GGTTCGACGAG 3252
 QY 405 PheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrVal 424
 Db 3253 TTCACCGTGACGGCGACGCGGCTCGGCTCTGCTCCAGGGCGGTG----- 3300
 QY 425 LysProAlaArgArgValAsnLeuProProValAlaValValSerProGlnLeuGlnGlu 444
 Db 3301 -----CCGGTGACCATCATCGATGGTCTCTCGCGGGG 3333
 QY 445 LeuThrLeuProLeuThrSerAlaLeuIleAsp----- 455
 Db 3334 AACTCTGCGCGCGCGCGTGAACCGGGCATCGATGCCAGCGGAAACGCTGGTGGTGGTGG 3393
 QY 456 -----GlySerGlnSerThrAspThrGluIleValSerTyrHisTrrpGluGlu 472
 Db 3394 ACGCTCAGCGGCTCCGCGAGCGACCGCGAGGGTGAACGCTGACGTACCATCCACTGGGAGCAG 3453
 QY 473 IleAsnGlyProPheIleGluGlyThrSerValAspSerProValLeuArgLeuSer 492

Db 3454 ATCGGTGGC-----ACGGATGTGACCTGACGGCGCGACACCTGACGCCGTGC 3504
Qy 493 AsnLeuAppProGlyAsnTySer-----PheArgLeuThrValThrAsp 507
Db 3505 TTCACCGCGCGAGCAGCGCGGTACACGCTCGGCTTTCATCTGACGGTCAGCGAT 3564
Qy 508 SerAspGlyAlaThrAsnSerThrAlaAlaLeuLeuVal-----AsnAsn 523
Db 3565 GGCACCTCGACACAGCGAGTCCGCGCTCATCGTGGCGGGAGCCCGGGTGCCCAAC 3624
Qy 524 AlaValAspTyProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGln 543
Db 3625 CCGAGCAACACTGCTCCGGAAGCGAGCGCGCGAGTCG-----GCCATCGTCGGCGAGGCG 3681
Qy 544 AsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleValLeuTyrglu 563
Db 3682 GCCACTGTGACGCTCAACGGCAGCGGACCGACGCTGACGGTGACACGCTGCTATCGTC 3741
Qy 564 TrpSer---LeuGlyProGlySerGluGlyLysHisValValMetGlnGlyValGln--- 581
Db 3742 TGGACGCGAGATCGGC-----GGCAGCGCGCGTCTGACGCGTCTCTG 3789
Qy 582 ---ThrProTyLeuHisLeuSerAlaMetGlnGluGlyAspTyThrPheGlnLeuLys 600
Db 3790 CTGACGCGCGAGTTCACCGCGCGCGCTCTCCCGAT---CCGCTGACCTTCTCTCTGATG 3846
Qy 601 ValThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnProGlu 620
Db 3847 GTACGCGACGCTAGC-----GCGACTGCTGACGAGTGTGATCCATCAGCGTGACCGAG 3900
Qy 621 AsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProValGlu 640
Db 3901 GAGAACGTGCGCGCGGTGCGCT-----CGCGCGGTGCTCTCTGCGCAACGACGACG 3954
Qy 641 SerAlaThrLeuAspGlySerSerSerAspAspHisGly---IleValPheTyHis 659
Db 3955 TCGCGCCAGCTCGACGCTCGGCTTCAGCGACGCGGACGCGATGCTGCTACGTACGCG 4014
Qy 660 TrpGluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAla 679
Db 4015 TGGACGCGAGGTCTCGCGCGCGAAGCGAC---ATCTCCGCGCGGACGACGCGGTGCG 4071
Qy 680 ThrValThr-----GlyLeuGlnValGlyThrTyHisPheArgLeuThrVal 695
Db 4072 GTGCTCAACCTGCTGACCTCGACGACAAAGCGGTGCTTCTGTTCTCGCTGACGCTG 4131
Qy 696 LysAspGlnGlyLeuSerSerSerThrThrLeuThrValAlaValLysLysGluAsn 715
Db 4132 AAGGACTCCGCGGACCGAGAGCAGCACACCGCTGCGAGTTCTCCGCGCGCAACGCGGT 4191
Qy 716 AsnSer-----ProProArgAlaArgAla-GlyGlyArgHis 727
Db 4192 GACAGCGGTGCTGCTCCGCGACGCGCGCGCGGTCTCGCGTCCGCTCGGAT 4238

RESULT 5

US-09-902-540-1119
; Sequence 1119, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1119
; LENGTH: 16584
; TYPE: DNA

; ORGANISM: Myxococcus xanthus
US-09-902-540-1119
Alignment Scores:
Pred. No.: 1.56e-17 Length: 16584
Score: 317.00 Matches: 162
Percent Similarity: 38.50% Conservatism: 84
Best Local Similarity: 25.35% Mismatches: 232
Query Match: 5.68% Indels: 161
DB: 4 Gaps: 28
US-10-764-390-3 (1-1072) x US-09-902-540-1119 (1-16584)
Qy 184 AspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGlyAspSerPro 203
Db 6679 GACCGGTCCATCGCGCGCGGTGAACCTGGTGAGCCTCTACGGTTCTCGGTCGCGATCCGGAA 6738
Qy 204 AlaValPro-----AlaGluThrGlnGlnAspProGluLeuHisTyLeu 218
Db 6739 GGCAGCGCGGTGACCTTCACCTGGGCACAGACG---TCTGGCCTTGGCGTCACTCGGTGGCG 6795
Qy 219 AsnGluSerAlaSerThrPro-----AlaProLysLeuProGluArgSerValLeu 235
Db 6796 GGTCCCAACACCTTGAACCCGTCGTTACGCGCGCGGAGGTTACCGAGTCCACCGCACTG 6855
Qy 235 ----- 235
Db 6856 GTGTTTCAGCTGACCGTCTCCGACGGGTCAAGACGTCACGAGCTCGGTGACCGTGACG 6915
Qy 236 LeuProLeuProThrThrProSerSer-----GlyGluValLeuGluLys 250
Db 6916 GTGCGGTCTCCCAACATCCGCCACCGTGACCGGGCCTCGACGGCATGTCGAGGAG 6975
Qy 251 GluLysAlaSerGlnLeuGlnSerSerAsnSer----- 263
Db 6976 CGCGCTGAGTACACGCTGAGCGGTCTCCGACGATGCGGATGCAATGCTCCATCGATAC 7035
Qy 264 -----SerGlyLysGluValLeuMetProSerHisSerLeuProAla 278
Db 7036 CTGTGGAGCCAGGTCTCCGGTACGCGGTGCGGTGAAGGACTACACACCGCGCGCG 7095
Qy 279 SerLeuGluLeuSerSerValThrValGluLysSerProVal-----LeuThrValThr 296
Db 7096 ACGTTCATCGCGCGCGAGGTACCGTGTGATGCTGCTTCCGCTTCCGCTGACGGTACG 7155
Qy 297 ProGly-Ser-----ThrGluHisSerIleProThrProThrSerAlaAl 312
Db 7156 GACGGCATCGCCACGCTGACGACACGCTGACCGTACCGGTACCAACCGCGCGCG 7215
Qy 312 aProSerGluSerThr-----ProSerGluLeuProLysSerProThrThrAlaPr 329
Db 7216 CCCATCGTCTCCGACACGCTCGGTGCTTCCGCGCGGCACTGTACCGTACCGGCTCC 7275
Qy 329 oArgThrValLysGluLeuThr-----ValSerAlaGlyAspAsnLeuIleIle--Thr 346
Db 7276 GCGGTGACCGCGGATGTTGACGCGCTGACGTGAGTGGGAGCAGACGGGTGGTTCGACG 7335
Qy 347 LeuProAspAsnGluValGluLeuLysAla-----PheValAlaProAlaProVal 364
Db 7336 GTCCCATCATGCGCGCGGACACGCTCGGCTATCTCTTCGCCACCCCGGTCCCG----- 7389
Qy 365 GluThrThrTyAsnTyArgGluTrpAsnLeuIleSerHisProThrAspTyGlnGlyGlu 384
Db 7389 ----- 7389
Qy 385 IleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyVal 404
Db 7390 -----GGTTCGTACGAG 7401
Qy 405 PheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrVal 424
Db 7402 TTCACCGTGACGCGACGCGGCTCGGCTCTCGCTCCAGCGCGGTG----- 7449

Db 1435 GCGCCTTCTCCGGCAGCGTCTCCACCACAGCAGCAGCGCCACGCTGGAGTTCCCTCG 1494
Qy 197 -----SerSerValGly----- 200
Db 1495 GGCAACACCCCTGACCTTACGCTCGATGCGCGCGCGCACGCGCACCGCTCCCG 1554
Qy 201 -----AspSer 202
Db 1555 GTGAGTTGGCG 1614
Qy 203 ProAlaValProAla-----GluThrGlnGlnAspProGluLeuHisTyr----- 217
Db 1615 CCCTCGCTCCCG 1674
Qy 218 LeuAsnGluSerAlaSer-ThrProAlaProLysLeuProGluArgSerValLeuLeuPr 237
Db 1675 GTGCCCG 1734
Qy 237 oLeuProThrProSer-----SerGlyG1 246
Db 1735 ACCTGGCTCAACCCCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1794
Qy 246 uValLeuGluLysGluLysAlaSerGlnLeuGlnGlnSerSerSerSerSerGlyLy 266
Db 1795 TACATGACCG 1854
Qy 266 sGluValLeuMetProSerHisSerLeuProProAlaSerLeuGluLeuSerValTh 286
Db 1855 AAGGTGAATGAGACCGCGCG-----CCTTCTGTTTCAGCTTCGCTACCGC 1899
Qy 286 rValGluLysSerProValLeuThrValThrProGlySerThrGluHisSerIleProTh 306
Db 1900 CACTCGTTCGAGTCCG-----ACATCTACCCCGCGCGCT-----CCATC 1938
Qy 306 rProProThrSerAlaAlaProSerGluSerThrProSerGluLeuProIleSerProTh 326
Db 1939 TTCCCGTACTACGCGCGCGCGCTCTCGAGTTACCCACG-----ACGCGCTCGAGTGTAC 1995
Qy 326 rThrAlaPro-----ArgThrValLysG1 334
Db 1996 GACATCGCGCGATGGCGGTGCGCTCCATCTACTCCATCGACCGCGCGCGCTCCCG 2055
Qy 334 uLeuThrValSerAlaGlyAspAsnLeuIleThrLeuProAspAsnGluValGluLe 354
Db 2056 GGCTACATCGAGCAGCGCG----- 2074
Qy 354 uLysAlaPheValAlaProAlaPro-ProValGluThrThrTyrAsnTyrGluTrpAsnL 374
Db 2075 -----CCAACCGCTGATGTACCGCAACGC----- 2099
Qy 374 euLLeSerHisProThrAspTyrGlnGlyGluLysGlnGly----- 388
Db 2100 --GTGGGTGGCGCAGACGTCACGAGGAGAAACGCGAGCGGTTCGCGTTCGCGAGTGG 2157
Qy 389 --HisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrVal---PheLysV 407
Db 2158 TCCACCGCGCAG--CTGCGGTGGCGACATGTTCCGCGGCAAGACGTCAGTTCGCT 2215
Qy 407 alThrValSerSerGluAsnAlaPheGlyGlu-----GlyP 419
Db 2216 TCCGATTTGTACACACCGCGGTGGCGGTGCGGTTCGAGCTTCGACGACGAGTGGCT 2275
Qy 419 heValAsnValThrValLysPro---AlaArgArgVal----- 430
Db 2276 TCACCAAGTGGAGCGCATGCTTCAGCGCGCGGTGGACGAGCGCGCGCGCGCGCA 2335
Qy 431 -----AsnLeuProProValAlaVal-----SerProG 441
Db 2336 CCTGCAACCGCGCGGTGGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2395
Qy 441 lnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeu-----IleA 455
Db 2396 TGTATGACGCGGTGAGAAACCCGAGACGCGGGGCGACTGGAGCTCGTGGCCACACGCTGG 2455

Qy 455 spGlySerGlnSerThrAsp---AspThrGluLysValSerTyrHisTrpGluLysLeA 474
Db 2456 ACGCACCGCGCAGCTATACCGCGAGCGCCAGCCCTCACCTACCTACCTGAGCAGGTGA 2515
Qy 474 enGlyProPheIleGluLysThrSerValAspSerProValLeuArgLeuSerAsnL 494
Db 2516 GCGGG-----CCGCGCGTACCTTGGTGAACG 2542
Qy 494 euAspProGlyAsnTyrSerPheArgLeuThrValThrAspSer----- 508
Db 2543 CGGACACCGCGCACCCCGTCTGTTTACCGCGGATGTCCCGACGACTCCATCTCGAGTTCC 2602
Qy 509 -----AspGlyAlaThrAsnSerThr-----ThrAlaAlaLeuLys 521
Db 2603 AGCTCGTGGTGGCGATGGCTCGGAGCACCGCGCGCACACGAGTCTCGGCGCTGGGTCC 2662
Qy 521 alAsnAsnAlaValAspTyrProProValAlaAlaGlyProAsnHisThrIleThrL 541
Db 2663 TCACAC-----GTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2698
Qy 541 euPro-----GlnAsnSerIleThrLeuAsnGlyAsnGlnSerS 554
Db 2699 CGCGCGCGAGGTGGCGAGTTCTCTCGGAGCGCGTGACGTGACCGCGCGCTCGA 2758
Qy 554 erAsp---AspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyL 573
Db 2759 GCGACCGCGAGCGGAGCTCATCTCTACAGTGGCG-----CAGACGTCGCGTC 2809
Qy 573 ybHisValMetGln-----GlyValGlnThrProTyrLeuHisLeuSerAlaMetG 591
Db 2810 CCGCGCTCGCGCTCAAGCGGGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2869
Qy 591 ln---GluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerT 610
Db 2870 CCGTCGATCTACGTTTCCCTTACCGTGGTGGTGAACCACTGCGAGGACCCAGAGCAAGA 2929
Qy 610 hrAla***ValThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaG 630
Db 2930 CGGCACCGTACGCTGATGTGGCG---AACGTGACCGCGCGCGCGCGCGCGCGCG 2986
Qy 630 lyProAspLysGluLeuPheProValGluSerAlaThrLeuAspGlySerSerSers 650
Db 2987 GC---TCCGACAGTGGTGGACCCGCTGGCGCACCGGAGCTCAGCGCGAGCTGTGTG 3043
Qy 650 erAspAspHisGlyIleValPheTyrHis-TripGluHisValArgGlyProSerAlaVal 669
Db 3044 GACCGGACCGGGATGCCCTCACCTGACGTGGCGTCAAGTGGAGGGCCCC---GCGGTG 3100
Qy 670 GluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnVal----- 686
Db 3101 GAGTCGAGGAGCGCACCGCGCACCGCGTTCACCGCGCGCGCGCGCGCGCGCGCGGAG 3160
Qy 687 GlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThr---Ser 705
Db 3161 ACGETGTGCGCTTCAG---GCCATCGCCAGCGCAACCGGCTGTCCAGCGCGCGCGGATG 3217
Qy 706 ThrLeuThrValAlaValLysLysGluAsnAsnSerPro-----ProArg 720
Db 3218 ACGCGAGTGTGCGGTGGCGCACCGCAACAGGGCGCGCGCGCGCGCGCGCGCGCG 3271

RESULT 7

US-09-248-796A-3204
; Sequence 3204, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Kelch Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 3204
 ; LENGTH: 2481
 ; TYPE: DNA
 ; ORGANISM: Candida albicans
 US-09-248-796A-3204

Alignment Scores:

Pred. No.: 1.73e-09 Length: 2481
 Score: 215.50 Matches: 198
 Percent Similarity: 32.80% Conservative: 91
 Best Local Similarity: 22.47% Mismatches: 355
 Query Match: 3.86% Indels: 239
 DB: 4 Gaps: 40

US-10-764-390-3 (1-1072) x US-09-248-796A-3204 (1-2481)

QY 11 LeuLeuLeuVal-ThrIleAlaGlyCysAlaArgLysGlnCysSerGluGlyArgTh 30
 DB 260 CTACTTCTACTCACAACAGCAGCCGGTGAACATCAAACTGTGAACACAAAATCAT 319
 QY 30 rTyrSerAenAlaValIleSerProAsnLeuGluThrThrArgIleMetArgValSerHi 50
 DB 320 CTCTCAGATCAATC-----ACCTCCACATCTCCCA 352
 QY 50 sThrPheProValAlaAspCysThrAlaAlaCysCysAspLeuSerSerCysAspLeuAl 70
 DB 353 CTCTTAC-----TGTTGCTGCTGCTGCAAC----- 380
 QY 70 atrPtpPheGluGlyArgCysTyrLeuValSerCysProHisLysGluAsnCysGluPr 90
 DB 381 -----TTCTTCAAGTCCTGTGTAGTACTCGTCCTGTT-CAACTTCGAACAAAAGCAA 435
 QY 90 oLysLysMetGlyProIleArgSerTyrLeuThrPheValLeuArgProValGlnArgPr 110
 DB 436 CAAGAAGAGAGTACAGCTAGACATCAACTT---CACCAGCTTACAACAGCTACTACTTCC 492
 QY 110 oAlaGlnLeuLeuAspTyrGlyAspMetMetLeuAenArgGlySerProSerGlyIleTr 130
 DB 493 AATACTCCTCTCTT----- 505
 QY 130 pGlyAspSerProGluAspIleArgLysAspLeu***PheLeuGlyLysAspTrpGlyLe 150
 DB 505 ----- 505
 QY 150 uGlnGluMetSerGluTyr***AspAspTyrArgGluLeuGluLysAspLeuGlnPr 170
 DB 506 -----CTCTTCCACC 516
 QY 170 oSerGlyLysGlnGluProArgGlySerAlaGluTyrThrAspTrpGlyLeuLeuProGl 190
 DB 517 TCCAGGAAACCCCAACCTCAAC-ACCTGCTCAG----- 549
 QY 190 ySerGluGlyAlaPheAsnSerSerValGlyAspSerProAlaValProAlaGluThrGl 210
 DB 550 -ACTAGTTCTGCCAACAAACAATCAACATCATCAATCAGCAGCTCCATCTACTAGTGT 608
 QY 210 nGlnAspPro-----GluLeuHisTyrLeuAsnGluSerAlaSer- 223
 DB 609 TATCCAAACCAAGTACATCAGAGTTCTATGTTCAATCACAACAACTTCTACAACACCAA 668
 QY 224 -----ThrProAlaProLysLeuProGluArgSerValLeuLeuPro----- 237
 DB 669 CACTCCAACACTAGTTCACCAATACTCCAACCACTTCAAGAAGCAGCAGCACTACTTTCAGC 728
 QY 238 -----LeuProThrThrProSerSerGlyGluValLeuGluLysGl 251
 DB 729 AGCACCACACATCAGAGGCTCCAGTTACTCCAAAGTACTTCGAAGAGTAGTACCAAAATAC 788
 QY 251 uLysAlaSerGlnLeuGlnSerSerAenSerSerGlyLysGluValLeuMetPr 271

DB 789 ACCAACCACA-----TCAAGAAGCACCACAAATACA-----CC 818
 QY 271 oSerHisSerLeuProAlaSerLeuLeuSerSerValThrValGluLysSerPr 291
 DB 819 AACCATATCAGAGGCTCCAGTTACTTCCAAAGTACTTCAAGTA---GTACCAAAATACACC 875
 QY 291 oValLeuThrValThrProGly-----SerThrGluHisSerIleProThrProProTh 309
 DB 876 AACCATTTCAAAAGCACCACAAATACACCAACCTCAGAACGACGACGATACACCAACCAC 935
 QY 309 rSerAlaAlaProSerGluSerThrProSerGluLeuProLysSerProThrAla-- 328
 DB 936 TTCAGAAGCACCACCAACTCCAAACCTCAGAGGCTCCGGTTACTTCAACACCACTCAGA 995
 QY 329 -----ProArgThr-----Vally 333
 DB 996 AGTAGTACCAACCACTTCAACTCAAGGAGAGTGTATCAACTTCATCATCACTAGTGTATC 1055
 QY 333 sGluLeuThrValSerAlaGlyAspAsnLeuIle-----IleThrLeuProAspAs 350
 DB 1056 CGACAGACAACTTTAACTTCTACTTCTATTAATCTCAACTCTGCTCTGACTACACA 1115
 QY 350 nGluValGluLeuLysAlaPheValAlaProAlaPro-----ProValGluThrTh 367
 DB 1116 GACTAGTACTCCAGAGGCTTCGATTTCTCTAAACCAAGTTCAACTTCTATTGAAACACC 1175
 QY 367 rTyrAsnTyrGluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGluLeuLysGl 387
 DB 1176 ATCAACTTCAACTTTGAACCAAGATCCAACTACTTCCAGTGTGGTACCCCTTCATC 1235
 QY 387 nGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysVa 407
 DB 1236 GGAGCAACTCAACCAACTACTTCTTGAAGTGGCT-----GT 1274
 QY 407 lThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAl 427
 DB 1275 AACATCAATTTCTCAACTCAAGAATCAACTTCTTGGTTGAACCAACCAACTAGTCTTT 1334
 QY 427 aArgArgValAsnLeuProValAlaValValSerProGlnLeuGlnGluLeuThrLe 447
 DB 1335 GGAAGTTTCCAACTCCA-----ACACCAACCCCTTCCACTTCTGAA--GCCCA 1382
 QY 447 uProLeuThrSerAlaLeuIleAspGlySerGlnSerThrAspAspThrGluIleValSe 467
 DB 1383 ACCATCAACTAGTGCA-----AGTCAAGCTCCACAGATACC----- 1419
 QY 467 rTyrHisTrpGluGluIleAsnGlyProPheIleGluLysThrSerValAspSerPr 487
 DB 1420 -----ACTTCTAGTGCACACGACACC 1439
 QY 487 oValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSerPheArg----- 502
 DB 1440 AGAATTTCTTCAAGTAAATGCAGATTTCTCCAACTCGTATTGTCATAGTTCGCAACACC 1499
 QY 503 -----LeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuI 520
 DB 1500 TTCATTAGTGAACCTTACCGATTCCAAATTTGATAGCTTCTCCACTACTGATGCC----- 1554
 QY 520 eValAsnAsnAlaValAspTyrProProValAlaAlaAsnAlaGlyProAsnHisThrIleTh 540
 DB 1555 -GTTTTCGCAAGCTACTACTGAACCGACTTCTGAAAACACA---CCTACTGCTGCTTCTTC 1610
 QY 540 rLeuProGlnAsnSerIleThrLeu-----AsnGlyAsnGlnSerSe 554
 DB 1611 TGTAAGTGCATGATGATTAATTCAGCCCAATATTTCAGCTCCAACTTCTTAATGTCAGCGC 1670
 QY 554 rAspAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHi 574
 DB 1671 CGAAACTGCTTCACTCCAGTTTCTGACAAAGTTTAGCCACTGGAAGTCAACACAGTCT 1730
 QY 574 sValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGlyLys 594

```

Db 1731 TGATACTACTGCTGGCGCTAGTTCAACTGCCAGTGAAGCAACTGCT-----GAGAAATT 1784
Qy 594 pTyThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla***ValTh 614
Db 1785 GTCTACTTTTGGT-----ACTGATGGCTCTTCAGATGCTTCT-----CA 1823
Qy 614 rValIleValGlnProGluAsnAsnArgPro-----ProValAlaVa 628
Db 1824 AACATTGGCGAAACCAACCAACCAATTCTCTGATCAATCACTAGTGGTTACACCTCAGCTAG 1883
Qy 628 lAlaGlyProAspLysGluLeuIlePheProVal-----GluSerAlaThr-----Le 644
Db 1884 TGCTTCTCCAGATGATATACACTC---CCAACTGGATCTGAAAGTGTGACACGCCCTGT 1940
Qy 644 uAspGlySerSerSerSerSerAspHisGlyIleValPheTyThrHisTrpGluHisValar 664
Db 1941 TTCTGGATCTGAAACCAAGTATTGATCAACCAATACAGTT-----GCTTC 1982
Qy 664 gGlyProSerAlaVal---GluMetGluAsnIle-----674
Db 1983 AGGATCTACTGTTATCCAGAAACAGTAATATTCCCAACCAATCACCATCACAATCTGT 2042
Qy 675 -----AspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyThrHisPh 691
Db 2043 AGTATCATCATGCTGCTGCTTCTTAACGCTCTCAACAGTTTCAGCAACCACTGATTCTCT 2102
Qy 691 eArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThr-----Va 709
Db 2103 TGCTGGAAGTGAACACCGAGTACACCAATATCATCAAGTCTACAGTACTAGTGAACC 2162
Qy 709 lAlaValLysGluAsnSerProArgAlaArgAlaGlyGlyArgHisValLe 729
Db 2163 AGTATTTTCAAGTCAATTAATTTCA-----AGTGAAGGTACCACTAGTTT 2207
Qy 729 uValLeuProAsnAsn-----SerIleThrLeuAspGlySerArgSerThrAspAs 746
Db 2208 GGTGTTCTTACCATTCTGAACTTAGCTCAACTGTTACTGTTAGTTCTGAACTGTGTC 2267
Qy 746 pGlnArgIleValSerTyThrLeuTrpIleArgAspGlyGlnSerProAlaAlaGlyAspVa 766
Db 2268 TACTGCTATTAAATCTGAATCGGTAACTGTTAGTTCTGATAGTCTGCT-----AC 2321
Qy 766 lIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyValTy 786
Db 2322 AGTTACTGGTAGTGAATCAATTTTGACTGGTGAATTAATCAAGTGAACAAAGTGTCTACTGCTATTGTC 2381
Qy 786 rThrPheHisLeuArgValThrAspSerGlnGlyValaSerAspThrAspThrAlaThrVa 806
Db 2382 AAGTGAATCTACTTTAACTGGATCCACTACTGTTGCTACTGCTACTGCTCCGCTGCACACTACTAT 2441
Qy 806 l 806
Db 2442 T 2442

```

RESULT 8

```

US-09-902-540-2872
; Sequence 2872, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 2872
; LENGTH: 1119
; TYPE: DNA

```

```

; ORGANISM: Myxococcus xanthus
US-09-902-540-2872

```

```

Alignment Scores:
Pred. No.: 2,34e-09 Length: 1119
Score: 207.50 Matches: 95
Percent Similarity: 41.40% Conservative: 59
Best Local Similarity: 25.54% Mismatches: 158
Query Match: 3.72% Indels: 60
DB: 4 Gaps: 17

```

```
US-10-764-390-3 (1-1072) x US-09-902-540-2872 (1-1119)
```

```

Qy 410 SerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgArg 429
Db 52 ACGGATGAGCGCGCGGGGCTTTGGCTGTCATCGATGACATCGCTTTCACGGGGTTG 111
Qy 430 ValAsnLeuProProValAlaValSerProGln-----441
Db 112 CGNACACGCCCTTCACGGCGCTGGTGACAGCAGCGGGGTCTGCACGACGGGCCATGTG 171
Qy 442 -----LeuGlnGluLeuThrLeuProLeuThrSerAla 452
Db 172 GATGCGCAGCGCGCGAGGACCTGGCGGTTCGATGAGCTGTCCGTG-----GTG 219
Qy 453 LeuIleAspGlySerGlnSerThrAspAspThrGluIleValSerTyThrTrpGluGlu 472
Db 220 CAACTGACGGCGACCGCGCAGCGCTGGCGGTGTGCTCAGCTACCGCTGGGAGCGAG 279
Qy 473 IleAsnGlyProPheIle-----GluGluLysThrSerValAspSer 486
Db 280 GTGCGGGTCCCTTCGCGCGGTAGTGCGCGGACACGCCGACCGCCAGCTTCACCGCG 339
Qy 487 ProValLeuArgLeuSerAsnLeuAspProGlyAsnTyThrSerPheArgLeuThrValThr 506
Db 340 CCGGAGGTGTGGCGCATACCATG-----CTGACCTTCGCTGACGGTGTG 387
Qy 507 AspSerAspGlyAlaThrAsnSerThrAlaAlaLeuIleValAsnAsnAlaValAsp 526
Db 388 GACGCGC---GGTTCGCGGACAGCGACACGCTCAGGTGCTCGTGCGCCAG---GTGAAC 441
Qy 527 TyrProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIle 546
Db 442 AAGCGCGCGCGGTGTCCGCGGACCCGCGCAGACGGTGAC---GAGGGGAGCACCGGTG 498
Qy 547 ThrLeuAsnGlyAsn---GlnSerSerAspHisGlnIleValLeuTyThrTrpSer 565
Db 499 ACGCTCCAGGGAAGCGCGGAGGACCGCGACCGCGCGGTGTGTCAGCCAGCAGTGGACG 558
Qy 566 LeuGlyProGlySerGluGlyLysHisValValMetGlnGlyValGln-----ThrPro 583
Db 559 -----CAGGTCTCCGCGCTCCGCTCAGCTGACGGCGCTGTATACGCTGACCCCG 609
Qy 584 TyrLeuHisLeuSerAlaMetGlnGluGly-----AspTyThrPheThrLeuLysVal 601
Db 610 AGCTTTCACCGCGCGCGGTGCGCAAGGAGCGCGGAGCTGTCTTCCAGCTCGTGTG 669
Qy 602 ThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnProGluAsn 621
Db 670 AGTGACGGGCAGTTGAGTGGCGGAGCGCGCGCGGTGAGTGTGACGGTGGCG-----720
Qy 622 AsnArgProProValAla-----ValAlaGlyProAspLysGluLeuIlePheProVal 639
Db 721 ---CACGTGCGCTGCGCGCCACCGTGTAGCGCGGGGACGATGTGGCCACCTTCTCGCG 777
Qy 640 GluSerAlaThrLeuAspGlySerSerSerSerAspHisGlyIleValPheTyThrHis 659
Db 778 AAGTCCGTCACGTGACCGCGCACCGGACCGGACCGGACGAGGAGCGGACGCTACCGCG 837
Qy 660 TrpGluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAla 679
Db 838 TGGAGCCAGACGCGGGGCCCCACC---GTGACGCTCCAGAACCGGACGCGCGCGGTGTC 894

```



```
QY 680 ThrValThrGlyLeuGlnValGlyThr-----TyrHisPheArgLeuThrVallys 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 895 CAGTTCACCGCGCGCTGAGTCTCAGCGCAACGGAGTTGCTTCTGTTGTAAGTCTCC 954
QY 697 AspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAlaVallys-----Ly 713
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 955 TCGGAGAGTGGCTCTCGCGGAGGACCGGTTCAAGGTGACGGTGAGCGCCGCGCCCG 1014
QY 713 sGluAsnAsnSerProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAs 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1015 GCGCGAAGGAGTCCAGCGGTGCTCGTCCACCGGACGA-----CTCCCGCG 1062
QY 733 nAsnSerIleThrLeuAspGlySerArgSerThr 744
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1063 CTCCCGCTGGCGATGCGGTTGCTCTCGCTCTCTCG 1096

RESULT 9
US-09-540-540-1013/c
; Sequence 1013, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIORITY FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIORITY FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1013
; LENGTH: 10838
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-540-540-1013

Alignment Scores:
Pred. No.: 1,37e-07 Length: 10838
Score: 207.50 Matches: 95
Percent Similarity: 41.40% Conservative: 59
Best Local Similarity: 25.54% Mismatches: 158
Query Match: 3.72% Indels: 60
DB: 4 Gaps: 17

US-10-764-390-3 (1-1072) x US-09-902-540-1013 (1-10838)
QY 410 SerSerGluAsnAlaPheGlyGlyGluPheValAsnValThrVallysProAlaArgArg 429
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10787 ACGGATGAGCGCGCGGGGCTTGGCTGTTTCATCGATGACATCGCTTACCGGGTTG 10728
QY 430 ValAsnLeuProProValAlaValSerProGln----- 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10727 GCGAACACACGCTCTTCCAGCGGCTGTGTGACGACGACGAGCGGGTCTGACACGCGGCCCATGTG 10668
QY 442 -----LeuGlnGluLeuThrLeuProLeuThrSerAla 452
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10667 GATGCGCAGCGCGCGGACGACCTGCGGTCGATGAGCTGTCGGTG-----GTG 10620
QY 453 LeuIleAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHisTrpGluGlu 472
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10619 CAACTGACGCGCAGCGCAGCAGCTGCGGGGTGCGCTACGTAACCGTGGGAGCAG 10560
QY 473 IleAsnGlyProPheIle-----GluGlyLysThrSerValAspSer 486
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10559 GTGGCGGTCCTTCGCGCGGATGAGTGGCGCGGACACAGCGCCACGCTTACCGCG 10500
QY 487 ProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSerPheArgLeuThrValThr 506
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10499 CCGGAGGTGTGCGCCGATACCATG-----CTGACCTTCGCTGACGGTGTGCTG 10452
QY 507 AspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuIleValAsnAsnAlaValAsp 526
```

```
Db 10451 GACGCG---GGTCTGCGGGACAGGGACACCGGTCCAGGTCTCGTGCGCCAG---GTGAAC 10398
QY 527 TyrProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIle 546
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10397 AAGCGCGCGCGGTGTCGCGGGCACCCGCGCAGACGGTGGAC-----GAGGGGAGCAGGTG 10341
QY 547 ThrLeuAsnGlyAen---GlnSerSerAspHisGlnIleValLeuLeuTyrGluTrpSer 565
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10340 ACGTCCAGGGAACGCGGAGGACCCGCGCAGCGGTGGTTCAGCACGACGATGGACG 10281
QY 566 LeuGlyProGlySerGluGlyLysHisValValMetGlnGlyValGln-----ThrPro 583
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10280 -----CAGGTCTCGCGCTCCGCTCAGCTGACGGCGCTGATACGCTGACCCCG 10230
QY 584 TyrLeuHisLeuSerAlaMetGlnGluGly-----AspTyrThrPheGlnLeuLysVal 601
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10229 AGCTTACCGCGCGCGGTGCGCCACAGGGGAGCGGAGCTGCTTCCAGCTCGTGTG 10170
QY 602 ThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnProGluAsn 621
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10169 AGTACGGGAGTTGAGTGGCGGAGCGCGGACGCGTGTGAGTGTGACGGTGC 10119
QY 622 AsnArgProProValAla-----ValAlaGlyProAspHisGlyIleValPheTyrHis 639
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10118 ---CACGTGCGCTCGCGCCACCGTACGCGCGGGGACGATGTGGCCACCTTCTCGCG 10062
QY 640 GluSerAlaThrLeuAspGlySerSerSerSerAspHisGlyIleValPheTyrHis 659
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10061 AAGTCCGTACGCTGACGCGCCACCGGACGCGCGGAGCGGAGCGCTGACCTACCG 10002
QY 660 TrpGluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAla 679
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10001 TGGAGCCAGACGAGGGGCCCCACC---GTGACGCTCCAGAACGCGCAGCGCGGTGTC 9945
QY 680 ThrValThrGlyLeuGlnValGlyThr-----TyrHisPheArgLeuThrVallys 696
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9944 CAGTTACCGCGCTCGAGGTCTCAGCGCCAACAGGAGTTCTGCTGTGTGAAGTCTCC 9885
QY 697 AspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAlaValLys-----Ly 713
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9884 TCGGAGAGTGGCTGTCCCGCAGAGACCGGGTCACGGTACGGTGAGCCCGCGGCCCG 9825
QY 713 sGluAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAs 733
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9824 GCGCGAAGGAGTCCAGCGGTGCTGCTGCCACCGGACGA-----CTGCCGCG 9777

RESULT 10
US-09-453-702B-39
; Sequence 39, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-DEC-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296.95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 25165

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 39:

US-09-453-702B-39

Alignment Scores:

Pred. No.: 2,27e-06 Length: 25165

Score: 201.50 Matches: 157

Percent Similarity: 34.40% Conservative: 102

Best Local Similarity: 20.85% Mismatches: 283

Query Match: 3.61% Indels: 211

DB: 3 Gaps: 30

US-10-764-390-3 (1-1072) x US-09-453-702B-39 (1-25165)

```

QY 190 GlySerGluGlyAlaPheAsnSerSerValGlyAspSerProAlaValProAlaGluThr 209
DB 14920 GGATCTGATGCGCTT---TCTGGCACCAGCGATCAGCTCGCGGTACGCGGATCACC 14976
QY 210 GlnGlnAspProGluLeuHisTyrLeuAsnGluSerAlaSerThrProAlaProLysLeu 229
DB 14977 -----GTCACCTGACGGGACAAACTACAGCGCACCGCATCGATGCTCC----- 15021
QY 230 ProGluArgSerValLeuLeuProLeuProThrProSerSerGlyGluValLeuGlu 249
DB 15022 GCAACTGGAGTGTACCGTCTGCTCGCGGGTCAGTGGCTGGCGGAAGCGACCTAC 15081
QY 250 LysGluLysAlaSerGlnLeuGlnGlnSerSerAsnSerSerGlyLysGluValLeu 269
DB 15082 AGCGTGACGGCGGCGGTACCAATGCTCAGGGTAAACAGCAGCACCGCCAGCCATAACGTG 15141
QY 270 MetProSerHisSerLeuProProAlaSerLeuGluLeuSerSerValThrValGluLys 289
DB 15142 CAGGTTAATACCGCTGCGGGCATCACCATTAAT-----CCGGTGGCGAGG 15189
QY 290 SerProValLeuThrValThrProGlySerThrGluHisSerIleProThrProThr 309
DB 15190 GACGATATTATCAACGCTTCAGAGCGAGCAGCGCGCAACTATCAGCGCCAGGTAC 15249
QY 310 SerAlaAlaProSerGluSerThrProSerGluLeuProIleSerProThrThrAlaPro 329
DB 15250 GGTGGCGCGCGGAGCAGCAGCGTTCACCGTTGAACCTGGGCGGTAAACTTACACCGCC--- 15306
QY 330 ArgThrValLysGluLeuThrValSerAlaGlyAspAsnLeuIleThrLeuPro--- 348
DB 15307 -----ACCGTCCAGGCTGATTTAAGCTGGAATGTGAGCGTACCTGCC 15348
QY 349 -----AspAsnGluValGluLeuLysAlaPheValAla----- 359
DB 15349 GCCGACTGCGAGGCTCTGGGTAAACGGTGAGTTAAACGGTTAATGCTCGGTGACCAACGCC 15408
QY 360 -----ProAlaPro----- 362

```

```

DB 15409 GTTGGCAACACCGGTAGCGGACCGCGGATATCACCATCGATGTCCTCGGCGCTG 15468
QY 363 ProValGluThrThrTyrAsnTyrGlu---TrpAsnLeuIleSerHisProThrAspTyr 381
DB 15469 CGGGTGGATACCGTCCGGGGTATGATGTGTGTCATATCATCGAACACGCT----- 15519
QY 382 GlnGlyGluIleLysGlnGly----- 388
DB 15520 CAGCGCAGGTGATCACCAGCAGCGAGTCCGGCTTTCGGCAGGACCGCGTTGACGGTG 15579
QY 389 -----HisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLys 406
DB 15580 GTGATTAAACAACTTACCGCCGACGCGTGTGGCGAAGCGTAGCTGGAGCGTCGCG 15639
QY 407 ValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysPro 426
DB 15640 GTTCTCGCGAGGATGTCACTAGTAACTGCGGCGGACGCTGAATATTACCGTTAGCGGG 15699
QY 427 AlaArgValAsnLeuProProValAlaValValSerProGlnLeuGlnGluLeuThr 446
DB 15700 GCGAACAGTCCGGA---ACGCAAAACGAGCATTAACCCATCCG-----CTGACT 15744
QY 447 LeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThrAspAsp----- 462
DB 15745 GTTGATCTCACCAGCGTCCCATCAGCATGAACAGCATCACAGCGATGATCGGATTAAC 15804
QY 462 ----- 462
DB 15805 GCCGCGAAAAAGCGCGGCTTAACGCTCTCCGCGCAGCAGCTCCGCGGTGGAACGGGG 15864
QY 463 ---ThrGluIleValSerTyrHisTrpGlu-----GluIleAsnGly 475
DB 15865 CAAACCGTCAACCGTCACTTTGGCGGCAAACTACACACGCGTGGCGGCAACCGGT 15924
QY 476 ProPheIleGluGluLysThrSerValAspSerProValLeuArgLeuSerAsnLeuAsp 495
DB 15925 AGCTGG-----AGCACACCGTTCGGCGGGGATTTGGCGGCTCGGT 15969
QY 496 ProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThr 515
DB 15970 GATGCGATGCCAGCGCCAGGTGGGTGACTTAACGTCAACGGC-----AACAGCGCC 16023
QY 516 ThrAlaAla-----LeuIleValAsnAlaVal 525
DB 16024 ACAGCAACGACGAATACAGCGTCGATAGCCCGCCGCAACCGTGACCATCAACCACT 16083
QY 526 AspTyrProProValAlaAsnAla-----GlyPro 535
DB 16084 GCCAGCGTAACTATCATCAGCCAGCAGCGGCGGGCGGTAAACGGTGTCCGCGACC 16143
QY 536 AsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAsp 555
DB 16144 AGCACCGCGAAACCGGCGCAGCGTCCCGTACCGTTCAGCTTAACGGCACT----- 16191
QY 556 AspHisGlnIleValLeuTyr-----GluTrpSerLeuGlyProGlySerGluGly 572
DB 16192 AACTACCAAGACCGGTGCAGACAGCGCAGCTGGAGTTA----- 16233
QY 573 LysHisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGlu 592
DB 16234 -----ACGTGCGCGCAGCGACTTACCGCACTTCTGCTAAT 16269
QY 593 GlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla*** 612
DB 16270 AACGCTACACCTCAGCCCGCAGCGTACGCGTCTGGCGGTAACTTTGGCGAGCGCCAGC 16329
QY 613 ValThrValIleValGlnProGluAsnAsnArgProProValAla-----ValAla 629
DB 16330 AAAGCGGTGACCGTC-----GATACCACTGCGCGGTGAGCTTTTAAACACCGTGGCG 16383
QY 630 GlyProAsp-----LysGluLeuIlePheProValGluSerAlaThrLeu 644

```

```
Db 16384 GCGCATGATGTGATTAAACACGTCGAACACATTTCAGCGCGCAGATTCATCAGCGCGCCCGCC 16443
QY 645 AspGlySerSerSerSerAspAspHisGlyIleValPheTyrHisTrpGluHisValArg 664
Db 16444 ACGGGCGCGTGGCGGGCAGCCCGCTGGTGGTGCACCATCGCGGCGCAGCATGTATGTACCC 16503
QY 665 GlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeu 684
Db 16504 AGCACCGATGCCAGCGGCAACTGGAGCGTCGGCTGCCCTGCCAGCGTGATTTCCGGCCTT 16563
QY 685 GlnValGlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThr 704
Db 16564 GCGGATGGCAGCGTGACCATCAGCGCCACCATTCATCGCATAGCGCAGCAACAGCAGCAG 16623
QY 705 SerThr-----LeuThrValAlaValLysLys 713
Db 16624 CAGACCCACAGCATCAGGTGAACAGCGCGGCGAGTGCCTCTCGGTGACGACTATTCAGC 16683
QY 714 GluAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsn 733
Db 16684 GCGGATAACCTTATTAAACGCCCGCGAAGCGGCG----- 16716
QY 734 AsnSerIleThrIleuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrLeu 753
Db 16717 AGTGGCGTGCACCTGAGCGGCACCT----- 16740
QY 754 TrpIleArgAspGlyGlnSerProAlaAlaGlyAspValIle-----AspGly 769
Db 16741 -----GGCCTAATTTCGCGACAGGTACAGTGGTGCCTGTTGCTTACGCGC 16788
QY 770 SerAspHisSerValAlaLeuGln----- 777
Db 16789 AAAGGCTACAGCGCCACCATTCAGACCAACGGGAGCTGGAGCGTGAACGTGCTCGCGCG 16848
QY 778 ----LeuThrAsnLeuValGluGlyVal----TyrThrPheHisLeuArgValThrAspSer 795
Db 16849 GATGTTGCGGCACCTAGTGTATGCGACAGGTACAGCTACCGTTAGCGCCTCCGCTCAGGACAGT 16908
QY 796 GlnGlyAlaSerAspThrAspThrAlaThrValGluVal 808
Db 16909 GCGGAAACAGCGCCACCGCCTCGCGCAGCGTGGCGGTG 16947

RESULT 11
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)

; LOCATION: (84808)..(84808)
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600592)..(600592)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
```

```

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664854)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

US-08-916-421B-1
Alignment Scores:
Pred. NO.: 0.00462 Length: 1664976
Score: 201.00 Matches: 128
Percent Similarity: 35.63% Conservative: 74
Best Local Similarity: 22.57% Mismatches: 217
Query Match: 3.60% Indels: 149
DB: 4 Gaps: 26

US-10-764-390-3 (1-1072) x US-08-916-421B-1 (1-1664976)
Qy 298 GlySerThrGluHisSerIleProThrProThrSerAlaAlaProSerGluSerThr 317
Db 33829 GGAATATAGAGTTGGAAATTC---CCCAATTGGAGTTGGAGGATACAGAGTTAT 33773
Qy 318 ProSerGluLeuProIleSerProThrAlaProArgThrValLysGluLeuThrVal 337
Db 33772 CCT---CAAAATTTATGCTTCTCTACAATAGCCCT---ACTTACTCCAATAAATATA 33719
Qy 338 SerAlaGlyAspAsnLeuIleThrLeuProAspAsnGluValGluLeuLysAlaPhe 357
Db 33718 ACA-----ATA 33713
Qy 358 ValAlaProAlaProProValGluThrThrTyrAsnTyrGluTyrAsnLeuIleSerHis 377
Db 33712 GTAGAGCCGGCA-----TCGTGGACTTGGTATTCGCATCATATA----- 33674
Qy 378 ProThrAspTyrGlnGlyGlu-----IleLys---GlnGlyHisLysGln 391
Db 33673 -----TATTTGGGAGATGGTGGCAGTGTTTGGATAAAACCCAAAAATGCCATAT 33623
Qy 392 ThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSer 411
Db 33622 ACTTTCACATACATACACTTCAGAAAGGGTTTATCCAAATATATGAAAGTTTATACT 33563
Qy 412 GluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgValAsn 431
Db 33562 GCAGAAAATAGAAAACAGTATATGTAATAGACATAAAACCCCAATAGCAACTCTAT 33503
Qy 432 LeuProProValAlaValValSerPro-GlnLeuGlnGluLeuThrLeuProLeu----- 449
Db 33502 ATTTATCCA-----AATCTCGAAGTTATAAGATACAAATTCCTTTAGTCCA 33455
Qy 450 -----ThrSerAlaLeuIleAspGlySerGlnSerThrAspSerThrGlu-- 464
Db 33454 ATAAATAGTTATGACCCAGATGCTAATAGAAATGATTCAGAGTATATAATTACTACGGCACC 33395
Qy 465 -----IleValSerTyrHisTrpGluGlu----- 472
Db 33394 TTAATAGGTATTATCCTATATCTCCAAACTCACCCCATGGCAAGATATATGGATTAAAC 33335
Qy 473 -----IleAsnGlyProPheIleGluGlyLysThrSe 483
Db 33334 TTAACAGTTTATGATTCTTAATGGAAATGTTGCATGGAAAT-TACTCATCAATGAATTAAAC 33276
Qy 483 rValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSerPheArgLe 503
Db 33275 TATTATATCC-----AAAAGTTTCTTATAGGGAATTACACTGCTAAATTT 33231
Qy 503 uThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuIleValAsnAs 523
Db 33230 GGTAGTTTGGGATGGAATGGGAGGAGTTAATTCAACAACCTGTTAAGTTTAGTGTG----- 33176
Qy 523 nAlaValAspTyrProProValAla-----AsnAlaGlyProAs 536
Db 33175 -----ATAAATAGACCTCTGTGGCTCAATTTATATATCTCCAGATAAACCTGAACCTTA 33120
Qy 536 nHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAs 556
Db 33119 TGAGGATCTAGAAATTCGTTTCCAGTCTTATGACCCGGAGGGT----- 33077
Qy 556 pHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHisValva 576

```

```
Db 33076 ----GAAATGTCATATTACATTGGAATTTTGAGATGGGACC-----GT 33036
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 1664976
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:

Qy 576 lMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGlyAspTyrTh 596
;
Db 33035 TATCAATACACAGATACG---ATAGTTCTACAAATATGAAGACACGCGTACTATAC 32979
;
Qy 596 rPheGlnLeuLysValThrAspSerArgGlnGlnSerThrAla***ValThrValIl 616
;
Db 32978 AGTTACCTTAACAGTATTGTGACAAATATATACCAAGTAGTTCATATCAAAACAATAGT 32919
;
Qy 616 eValGlnProGluAenAsnArgProProValAlaValAlaGlyProAspLysGluLeuIl 636
;
Db 32918 AGTTGGGGT-----ATATTAGCTAACTTTTACATATGAATAAT 32880
;
Qy 636 e-----PheProValGluSerAlaThrLeuAspGlySerSerSe 650
;
Db 32879 AAATGAAATATCTGTAAGATTTTGAAGATACATCGTTGGCCCTCGGAA-- 32828
;
Qy 650 rAspAspHisGlyIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValI 670
;
Db 32827 -----ATAATTAGTTGGCCTGGGACITTTGGAGATGGCAGCACATCAACAGA 32781
;
Qy 670 uMetGluAenIleAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHi 690
;
Db 32780 ACAAATCTCATCCAT-----ACTTACGCAAGAGGGGGCATACTT 32739
;
Qy 690 sPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrLeuThrValAl 710
;
Db 32738 TGTAACTTAACTGTTTAAAGCGGATACAAACTTAACGGATAGTGAACAAAGATTGTTAT 32679
;
Qy 710 aValLysLysGluAenAsnSerProProArgAlaArgAlaGlyArgHisValLeuVa 730
;
Db 32678 AATTTCTCACCTCCAAATATCTCCAGTTGCTGACTTACGTATAAA-----GT 32628
;
Qy 730 lLeuProAenAsnSerIleThrLeuAspGlySerArgSerThrAsp---AspGlnArgIl 749
;
Db 32627 AATAAATGGAACACTACTGTAGAGTTTAAACGCTAATTTATCATAGATGAAGATGGATATAT 32568
;
Qy 749 eValSerTyrLeuThrIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGl 769
;
Db 32567 TAAATATTATATCTGG-----GACTT 32547
;
Qy 769 ySerAspHisSerValAlaLeuGlnLeuThrAenLeuVal----- 782
;
Db 32546 TGGAGATGGAACACGACGAAATACACAAATCCAAATATATATACACAAATATAAAA 32487
;
Qy 783 -GluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAs 802
;
Db 32486 AGAGGGGGTTTATCCAGTTTACCTTGACAGTTATAGATAATGATAATCTTTTCAGATTCAAG 32427
;
Qy 802 pThrAlaThrValGluVal 808
;
Db 32426 AACTAAGTTAATAAATATA 32408
;
RESULT 12
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
```

```

; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

Alignment Scores:
Pred. No.: 0.00462 Length: 1664976
Score: 201.00 Matches: 128
Percent Similarity: 35.63% Conservative: 74
Best Local Similarity: 22.57% Mismatches: 217
Query Match: 3.60% Indels: 149
DB: 4 Gaps: 26

```

US-10-764-390-3 (1-1072) x US-09-692-570-1 (1-1664976)

```

Qy 298 GlySerThrGluHisSerIleProThrSerAlaAlaProSerGluSerThr 317
Db 33829 GGAATATAGAGGTTGGAATTCCA---CCACCAATTGGAGTTGGAGGATACAGCAGTTAT 33773
Qy 318 ProSerGluLeuProIleSerProThrThrAlaProArgThrValLysGluLeuThrVal 337
Db 33772 CCT---CAATTTATGCTTCTCCACATAGCCCT---ACTTACTCCATAAATATA 33719
Qy 338 SerAlaGlyAspAsnLeuIleIleThrIleuProAspAsnGluValGluLeuLysAlaPhe 357
Db 33718 ACA-----ATA 33713
Qy 358 ValAlaProAlaProProValGluThrThrTyrAsnTyrGluThrAsnLeuIleSerHis 377
Db 33712 GTAGAGCGGCA-----TCGTGGAGCTTGGTATTGGCATCATATA----- 33674
Qy 378 ProThrAspTyrGlnGlyGlu-----IleLys---GlnGlyHisLysGln 391
Db 33673 -----TATTCGGAGATGGTGGCAGTGTTGGATAAAACCCAAAAATGCCCATAT 33623
Qy 392 ThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSer 411
Db 33622 ACTTTCACATACATACACTTCAGAGGGGTTTATCCAAATATATATGAAAGTTTATACT 33563

```

QY 412 GluAsnAlaPheGlyGluGlyPheValAsnValThrValValProAlaAlaArgValAsn 431
 Db 33562 GCAGAAATATGAACACAGTATATGTAATAGACAAATAAAACCAATAGCCAACTCTAT 33503
 QY 432 LeuProProValAlaValValSerPro-GlnLeuGlnGluLeuThrLeuProLeu----- 449
 Db 33502 ATTTATCCA-----AATCCTGCAAGTTATAAAGATACAAATTTCTTTAGTCCA 33455
 QY 450 -----ThrSerAlaLeuLeuAspGlySerGlnSerThrAspAspThrGlu-- 464
 Db 33454 ATAAATAGTTATGACCCAGATGCTAATAGAAAGATCCAGAGATATTAATTACTACGGCACC 33395
 QY 465 -----IleValSerTyrHisTyrGluGlu----- 472
 Db 33394 TTAATAGGTATATCTCTATATCTCCAAACTCACCCTGCGAAGATATATGATGATTTAAC 33335
 QY 473 -----IleAsnGlyProPheIleGluGluLysThrSe 483
 Db 33334 TTAACAGTTTATGATTTCTAATGGAATGTTGCATGGAAT-TACTCATCAATGAATTAAC 33276
 QY 483 rValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSerPheArgIle 503
 Db 33275 TATTATATCC-----AAAAAGTTTCTCTATAGGAATTAACACTGCTAAAT 33231
 QY 503 uThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuIleValAsnAs 523
 Db 33230 GGTAGTTTGGATGGATGGAGGAGTAAATTCACCAACTGTTAGTTTGTG----- 33176
 QY 523 nAlaValAspTyrProProValAla-----AsnAlaGlyProAs 536
 Db 33175 ----ATAAATAGACCTCTCTGTGGCTCAATTTATATATCTATCCAGATAAAACCTGAACCTAA 33120
 QY 536 nHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAs 556
 Db 33119 TGAGGATGAGAAATCGTTTCCAGTCTTATGACCCGAGGGT----- 33077
 QY 556 pHisGlnIleValLeuTyrGluTyrSerLeuGlyProGlySerGluGlyLysHisValVa 576
 Db 33076 ----GAAATGTCATATTACATTTGGAATTTGGAGATGGACC-----GT 33036
 QY 576 lMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAspTyrTh 596
 Db 33035 TATCAATACAAACAGATACG---ATAGTTTCATCAAAATATGAAGACAGGCTCTATATAC 32979
 QY 596 rPheGlnLeuLysValThrAspSerSerArgGlnSerThrAla**ValThrValI 616
 Db 32978 AGTTACCTTAACAGATATTGACAAATATAACGCAAGTAGTTCAATATCAAAACAAATAGT 32919
 QY 616 eValGlnProGluAsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuI 636
 Db 32918 AGTTGGGGT-----ATATTAGCTTAATTTACATATGAATATAT 32880
 QY 636 e-----PheProValGluSerAlaThrLeuAspGlySerSerSe 650
 Db 32879 AAATGAAATPACTGTGAAGATTTGAAGATACATCAGTTGTGGCCCTCGGAAA----- 32828
 QY 650 rAspAspHisGlyIleValPheTyrHisTyrGluHisValArgGlyProSerAlaValGl 670
 Db 32827 -----ATAATTAGTTGGCACTGGGACTTTGGAGATGGCAGCATCAACAGA 32781
 QY 670 uMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHi 690
 Db 32780 ACAAATCTATTCAT-----ACTTACAGCAAGAGGGGCATCTT 32739
 QY 690 sPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAl 710
 Db 32738 TGTAACCTTAACCTGTTAAAGCGCATACAACTTAACGATAGTGTAAACAAAGATTGTTAT 32679
 QY 710 aValLysLysGluAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuVa 730
 Db 32678 AATTTCTCCACCTCCAAAATATCTCAGTTGCTGACTTTACGTATATAA-----GT 32628
 QY 730 lLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAsp---AspGlnArgI 749

Db 32627 AATAAATGGAACCTACTGTAGAGTTTAAACGCTAAATTTATCATACGATGAAGATGGATATAT 32568
 QY 749 eValSerTyrLeuTyrPheIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGl 769
 Db 32567 TAAATATTATATCTGG-----GACTT 32547
 QY 769 ySerAspHisSerValAlaLeuGlnLeuThrAsnLeuVal----- 782
 Db 32546 TGGAGATGGAACACCGCAATATACAAACAATCCAAATTTATATACACAATATAAAAA 32487
 QY 783 -GluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAs 802
 Db 32486 AGAGGGGTTTATCCAGTTACCTTCAGACAGTTATAGATAATGATAATCTTTTCAGATTCAAG 32427
 QY 802 pThrAlaThrValGluVal 808
 Db 32426 AACTAAGTTAATAATATA 32408
 RESULT 13
 US-09-621-976-13082
 ; Sequence 13082, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621.976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 13082
 ; LENGTH: 320
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-621-976-13082
 Alignment Scores:
 Pred. No.: 1.56e-09 Length: 320
 Score: 199.00 Matches: 40
 Percent Similarity: 86.96% Conservative: 0
 Best Local Similarity: 86.96% Mismatches: 6
 Query Match: 3.57% Indels: 0
 DB: 4 Gaps: 0
 US-10-764-390-3 (1-1072) x US-09-621-976-13082 (1-320)
 QY 503 LeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuIleValAsn 522
 Db 180 TTGACTGTAGTAGACTCTGTAGAGCTACCACTCTACTCTGCAAACTGCAGTGAAC 239
 QY 523 AsnAlaValAspTyrProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuPro 542
 Db 240 AAAGCTGTGGATTACCCCTGTGGCCAAACGAGGCCCAACCAAGTGATCATCCCTGCC 299
 QY 543 GlnAsnSerIleThrLeu 548
 Db 300 CAAAACCTCATCACCCCTC 317
 RESULT 14
 US-09-902-540-1136/c
 ; Sequence 1136, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540

Db 13638 GCGGTAATACGTGGTTCTGTCACACACAGGCGGAGGATCTGGATTACACCTGG 13697
 Qy 373 -----AsnLeuIleSerHisProThrAspTyrGlnGly 383
 Db 13698 CCGAAGATGTGACCGCGCTACATACCTGACGGTAGAAGCGCAT----- 13748
 Qy 384 GluIleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyr 403
 Db 13749 -----AAGCGGGAATAAGACGACGAGCGCTCGAT----- 13781
 Qy 404 ValPheLysValThrValSerGluAsnAlaPheGlyGluGlyPheValAsnValThr 423
 Db 13782 -----TTTACCATTTGATACC----- 13796
 Qy 424 ValLysProAlaArgValAsnLeuProProValAlaValValSerProGlnLeuGln 443
 Db 13797 -----CGGCTGCAAGCGCTACCATCGCTATG----- 13823
 Qy 444 GluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThrAspAspThr 463
 Db 13824 -----GATAGCAGGAGCATACA 13841
 Qy 464 GluIleValSerTyrHisTrpGluGluIleAsnGlyProPheIleGluGluLysThrSer 483
 Db 13842 GGTGCCATTGGCATCAT-----ATTACGAGC 13868
 Qy 484 ValAspSerProValLeuArgLeuSerAsnLeuAsp----- 495
 Db 13869 GTCAAAAGACCGGGCTTTACTATTGGCAATATTGACGCGATGCGCACTCGTCAATTTG 13928
 Qy 495 ----- 495
 Db 13929 CCGATCACACAGGCGGCAATAGCCAGGAGTGACCTAACCCAGTTGGAGGACAGTGG 13988
 Qy 496 -----ProGlyAsnTyrSerPheArgLeuThrValThr 506
 Db 13989 CGCTTTACCGCAGATGCTAGCTGGCGGACGCTAGCTATACGCTGACGGTAGAGTTAAG 14048
 Qy 507 AspSer-----AspGlyAlaThr 512
 Db 14049 GATAACGCGGAAACGTTTCGTGCTCAGTCCACGCGCTGGTGGTGACGGTGACGCAACACC 14108
 Qy 513 AsnSerThrAlaAlaLeuIleValAsnAsnAlaValAspTyrProProValAlaAsn 532
 Db 14109 AGCATTAATGATATTACATTCGTCATATGATGCTGCTGCTGATGACAACTCTAACTAAT 14168
 Qy 533 Ala--GlyProAsnHisThrIleThrLeuProGln-----AsnSerIleThrLeuAsn 549
 Db 14169 AGCACCCGTCGCGAGTTGAGATACACGCTGCGCGGATGTGAATTCGTGCAACTGAGC 14228
 Qy 550 GlyAsnGlnSerSerAspHisGlnIleValLeuTyrGluTrp---SerLeuGlyPro 568
 Db 14229 ATTGATGGGGCGCAAC-----TGGGTGAGCGCGGCGAG 14264
 Qy 569 GlySerGluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSer 588
 Db 14265 GGTATCGAAGG-----GTCTGGGGCTATACCTGGCCA-----ACG 14300
 Qy 589 AlaMetGlnGluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGln 608
 Db 14301 GATATGGCGGATGGAACACACACCTTAACCGCTGATCCCGACGAGCGGGCAATACG 14360
 Qy 609 SerThrAla***ValThrValIleValGlnProGluAsnAsnArgProProValAlaVal 628
 Db 14361 GCGAGCAACCGTTGAATTTTCATCGACACCGGTTGTCGACCGCCACCATTTGGCGTG 14420
 Qy 629 AlaGlyProAspLysGluIlePheProValGluSerAlaThrLeuAspGlySerSer 648
 Db 14421 -----GAT 14423
 Qy 649 SerSerAspAspHisGlyIleValPheTyrHisTrpGluHisValArgGlyProSerAla 668
 Db 14424 AGCAGGATGATACCGGTACGCTGGCGATGATATGACCAATATGACCGCCGCGACCTTT 14483

Qy 669 ValGluMetGluAsnIleAspLysAlaIleAlaThrValThr----- 682
 Db 14484 ATT---CTGCAGAATATCGATTTCGATGTTTCAACGTTACAGTCAGCGTCACGCATAAT 14540
 Qy 683 GlyLeuGlnVal----- 686
 Db 14541 GGAACGACAACTCTGTTTACTGCGACACAGGGGCTGGAGATGGAGCTTTACACCGCCA 14600
 Qy 687 -----GlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGlyLeu 701
 Db 14601 GCGCGGTGGCGGACGCTGATTATACGCTGACGCTGAGAGGATCGGCGGGAAT 14660
 Qy 702 SerSerThrSerThr---LeuThrValAlaValLysGlu--- 714
 Db 14661 ACGCGTCTGCTACGCGCTGACGCTGACGTGATGATGACGAAATAGCCATTGATCGTATT 14720
 Qy 715 -----AsnAsnSerProArgAlaArgAlaGlyGlyArgHisVal----- 728
 Db 14721 GAATTAGTCAACGATACGCGGCTCCCTGGCGATTAATGTGACAAACATGTGCGTCCGCGAG 14780
 Qy 729 -----LeuValLeuProAsnAsnSer-----IleThrLeuAspGlySerArg 742
 Db 14781 TTCCAGATCTCGGTACCGGATGATGT---GGAAGAGTCTCTCTGATGATTGACGCGGCGACG 14839
 Qy 743 SerThrAspAspGlnArgIleValSerTyrLeuTrpIleArgAspGlyGlnSerProAla 762
 Db 14840 ACC-----TGGGTACTGCAATCAAGATTCGACG 14869
 Qy 763 AlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsnLeuVal 782
 Db 14870 GCTGCG---ATTGGGATTACAC---TGGCCGCGGATATGCGA 14908
 Qy 783 GluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAsp 802
 Db 14909 GAGGACAGCATACCTGACCGTGAAGTGAATGACTGACGCTGGGGTAAATAAGATGACGAG 14968
 Qy 803 ThrAla-----ThrValGluValGlnProAspProArgLysSerGlyLeuValGluLeu 820
 Db 14969 ACGCTCAATTTCATCTATCGATATCACG-----TTGTTAACGCGCA 15007
 Qy 821 ThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAsp-----ThrLeu 837
 Db 15008 ACCATTGAGCTAGCGCGCGATCAGATACCGGACAGATAGAACGATAATCTGACCGAGC 15067
 Qy 838 ValArgGlnLeuAlaValLeuAsnValLeuAspSerAspIleLysValGlnLysIle 857
 Db 15068 GTCACTCAGCGGTATTGTTGTTGGGAGTATCGATAAAGATGTTCGA----- 15115
 Qy 858 ArgAlaHisSerAspLeuSerThrValIleValPheTyrValGlnSerArgProPhe 877
 Db 15116 -----CACGTGGAATTGAGT-----ATTGACATAACGCGACGTTT 15151
 Qy 878 LysValLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLysGluLys 897
 Db 15152 AAAACGCTGTACTCACCGAATCAGCGGACGCTGGCGGCTATTCGACCGGATTCGCTTTG 15211
 Qy 898 AlaAsp---PheLeuLeuPheLysValLeuArgValAspThrAlaGly 912
 Db 15212 GCGGACGGTAGCTACACATTCACCGTACGCGTAACAGATGTGGCAGGC 15259

Search completed: October 12, 2005, 16:36:49
 Job time : 3747 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 12, 2005, 10:31:17 ; Search time 6860 Seconds
(without alignments)
5948.236 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 5580

Sequence: 1 MAPPTGVLSLLLVITAGC.....VSMNGSIRNGASFSYCSKDR 1072

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Rgapop 6.0 , Rgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool p/US10764390/runat 12102005 110203 22340/app query.fasta_1.1223
-DB=EST -OFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCF=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10764390.@CGN 1 1 5180 @runat 12102005 110203 22340 -NCPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gsl.*
9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	4006.5	71.8	4828	3 AK051381	AK051381 Mus muscu
2	2420.5	43.4	3660	3 AK084668	AK084668 Mus muscu
3	2420.5	43.4	4141	3 AK043006	AK043006 Mus muscu
4	2389.5	42.8	3616	3 AF289597	AF289597 Homo sapi
5	2310	41.4	3067	3 AK031342	AK031342 Mus muscu
6	2256.5	40.4	3671	3 HSM805357	AL834315 Homo sapi
7	1982	35.5	2959	9 AY417695	AY417695 Mus muscu
8	1940.5	34.8	2962	9 AY417693	AY417693 Homo sapi
9	1635	29.3	2962	9 AY417694	AY417694 Pan trogl

10	1471	26.4	2189	3 AK049570	AK049570 Mus muscu
11	1312	23.5	842	1 AU130115	AU130115 AU130115
12	1051	18.8	710	6 CB722905	CB722905 UI-M-GH0-
13	1025	18.4	597	1 AL044313	AL044313 DKFZp4340
14	999	17.9	598	5 BP229849	BP229849 BP229849
15	994	17.8	842	7 CK463785	CK463785 934714 MA
16	922	16.5	750	7 CN527744	CN527744 UI-M-H00-
17	916	16.4	534	7 CR536629	CR536629 DKFZp459A
18	910	16.3	927	5 BQ897270	BQ897270 AGENCOURT
19	890	15.9	821	1 AU129983	AU129983 AU129983
20	877	15.7	814	1 AU131313	AU131313 AU131313
21	870	15.6	815	7 CR566736	CR566736 CR566736
22	862	15.4	610	6 CB519102	CB519102 UI-M-GH0-
23	849	15.2	875	6 CA790104	CA790104 AGENCOURT
24	845	15.1	772	1 AU132238	AU132238 AU132238
25	841.5	15.1	909	5 BU422326	BU422326 603019005
26	836	15.0	494	5 BQ267742	BQ267742 i.j9e06.Y
27	820	14.7	785	4 B1766063	B1766063 603053031
28	818	14.7	1056	4 BM547013	BM547013 AGENCOURT
29	816	14.6	685	6 CB527694	CB527694 UI-M-FY0-
30	812	14.6	727	5 BQ768962	BQ768962 UI-M-FC0-
31	800	14.3	1107	4 BM479855	BM479855 AGENCOURT
32	792	14.2	479	1 AL589494	AL589494 DKFZp451G
33	789.5	14.1	1040	2 BE311921	BE311921 601143520
34	784.5	14.1	949	4 BF981495	BF981495 602309331
35	784	14.1	653	4 BI390119	BI390119 P9PIC.BK0
36	763.5	13.7	805	4 BI159118	BI159118 602922880
37	761	13.6	663	7 CN365571	CN365571 170004245
38	757	13.6	655	7 CF899230	CF899230 A0300F03-
39	753	13.5	777	7 CR765174	CR765174 DKFZp468E
40	752	13.5	638	1 AL858156	AL858156 AL858156
41	750	13.4	964	4 BG172757	BG172757 602337184
42	749.5	13.4	968	2 BE617479	BE617479 601442136
43	738	13.2	829	7 CF579859	CF579859 AGENCOURT
44	737	13.2	616	7 CN704544	CN704544 E0489B01-
45	737	13.2	617	7 CF170107	CF170107 B0822F05-

ALIGNMENTS

AK051381 4828 bp mRNA linear HTC 03-APR-2004
Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
enriched library, clone:DI30043K22 product:hypothetical PKD domain
containing protein, full insert sequence.

ACCESSION AK051381.1 GI:26094501

VERSION AK051381.1

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus

ORGANISM

REFERENCE 1 Carninci, P. and Hayashizaki, Y.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Qy 351 GluValGluLeuLysAlaPheValAlaProProValGluThrThrTyrAsnTyr 370
Db 1335 GAACGACAGTGAAGGCTCTGTGAACACGCGCCCTCGCAGATACAACTTACTCTAT 1394
Qy 371 GluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGluIleLysGlnGlyHisLys 390
Db 1395 GAATGGAGTTAATGAGCCACCCAGTAGACTTCCAGGTAAATCAACAGAAACACAG 1454
Qy 391 GlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSer 410
Db 1455 CCGACTCTTCACTCTCTCAATATATCTGTGGAGCTCTATGCTTTCAGAGTGGCTTCT 1514
Qy 411 SerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgVal 430
Db 1515 AGTGAGAATGCATTGGAGAGGCTATGTCAATGTCAAGGTATGCGCAGAGTGC 1574
Qy 431 AsnLeuProProValAlaValValSerProGlnLeuGlnLeuThrLeuProLeuThr 450
Db 1575 AACCAGCCACTGTAGCTGTGTGTTCTCCCGACACAGAGCTCAGTGTGCTTGAAC 1634
Qy 451 SerAlaLeuIleAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHisTrp 470
Db 1635 TCAGCCCTGATCGATGCGCAGCAAGTACAGATGATCTGAAATAGTGAGTTACCACTGG 1694
Qy 471 GluGluIleAsnGlyProPheIleGluLysThrSerValAspSerProValLeuArg 490
Db 1695 GAAGAAGTCGACGGGCGCTTCTCTAGGAGAGAGTTTCCAGCTGACACCCCAATCTTACGA 1754
Qy 491 LeuSerAsnLeuAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGly 510
Db 1755 TTGTGCAACCTAGTTCCTGTGTAACCTTACCTTTAGTTGACCATCACAGACTCGGATGGC 1814
Qy 511 AlaThrAsnSerThrThrAlaAlaLeuIleValAsnAsnAlaValAspTyrProProVal 530
Db 1815 GCTACCAACTCCACCACTGCATCCCTGCTCATCGTGATGCGTGAGTACCCACCGCGT 1874
Qy 531 AlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGly 550
Db 1875 GCCAATGAGGCGCCAAATCAGACCATCACTTGTGCCCAACACCATCATCTCAATGGG 1934
Qy 551 AsnGlnSerSerAspAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySer 570
Db 1935 AACGACAGCAGTGTATGATCACAGATGTTCTCTATGATGTTGCGAGGCGCTGGCGGT 1994
Qy 571 GluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMet 590
Db 1995 GAGACGAAAGAGATGGTCAATCGAGGCGCAGACCCCATACCTTCACTTGTCTGAGCTG 2054
Qy 591 GlnGluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThr 610
Db 2055 CAGGAAGGAGAGTACACATTTTCAGCTGATGTTGACGGATTCCTCGGGACAGCAATCCACC 2114
Qy 611 Ala**ValThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaGly 630
Db 2115 GCTTTGGTGTGACTGTCCAGCTGAGAACACACAGGCTCCGGTGGCGGTGGCAGGC 2174
Qy 631 ProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSer 650
Db 2175 CCCGACAGGAGAGTGGTTTCCCGGTGACAGTGTCTACACTGGATGGAGCGGAGCAGT 2234
Qy 651 AspAspHisGlyIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValGlu 670
Db 2235 GACGATCACGGCATTTGTCTGATCATCTGCGGAGCACATCAGAGGCGCCGAGTGCAGTGGAG 2294
Qy 671 MetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHis 690
Db 2295 ATGAGAAATGGGACAAAGCCATAGCCACTGTACCGGCGCTTACGGTGGGTATCTTACCAC 2354
Qy 691 PheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAla 710
Db 2355 TTCCGGTTAACCGTGAGAGACAGCAGGCTCTGAGCAGCAGTCCACCTCAGTGTGGCT 2414
Qy 711 ValLysLysGluAsnAsnSerProProArgAlaArgAlaGlyArgHisValLeuVal 730

Db 2415 GTGAGAAGAGAAATTAACGCCCTCCAGAGCCAGGCTGGTGACACATGTTCTTATA 2474
Qy 731 LeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleVal 750
Db 2475 CTTCCCAATAATTCCTACTTCTGGATGGTTCAAGGCTCTACTGATGACCGAGGAATGTG 2534
Qy 751 SerTyrLeuTrpIleArgAspGlyGlnSerProAlaIaGlyAspValIleAspGlySer 770
Db 2535 TCCTATCTGTGGATCCGGATGCCAGAGTCCAGCTCCCGAAC----- 2578
Qy 771 AspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeu 790
Db 2578 ----- 2578
Qy 791 ArgValThrAspSerGlnGlyAlaSerAspThrAlaThrValGluValGlnPro 810
Db 2578 ----- 2578
Qy 811 AspProArgLysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThr 830
Db 2579 ---CCCAAGAGAGATGGCCTGGTGGAGTTGATCCTCGAGGTGGTGTGAGCAGCTGACA 2635
Qy 831 GluGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSer 850
Db 2636 GAGCAGCAGAGAGAGACTCTAGTGAGCAGCTGGCTGTCTCTCGAACGTTGCTAGACTCA 2695
Qy 851 AspIleLysValGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTyr 870
Db 2696 GACGTGAAGGTGTTGAAGATCCAGGCTCACACGGATGTCAGCACTGTGATTGTTATTAT 2755
Qy 871 ValGlnSerArgProPheLysValLeuLysAlaAlaGluValAlaArgAsnLeuHis 890
Db 2756 GTACAGAGTGGATCACCTTTCAAGGTTCTCAGAGCTGCTGCTGTGGCCCGGAATCTGCAC 2815
Qy 891 MetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAspThr 910
Db 2816 AAGCGGCTTTCCAAAGGAGAGAGGCTTTCTGCTTTTCAAGGCTTGTGAGGTTAGACACA 2875
Qy 911 AlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCys 930
Db 2876 GCAGGGTGCTTCTCAAGTGTTCAGGCCACGGTCACTGTGACCCCATCACAGCGCTGC 2935
Qy 931 IleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyGlnSer 950
Db 2936 ATCTGCTCCAGCTGTGGATGGAGAACCTCATACAGCGGTACATGTGGGATGGAGAGAGC 2995
Qy 951 AsnCysGluTrpSerIlePheTyrValThrValLeuAlaPheThrLeuIleValLeuThr 970
Db 2996 AACTGTGAGTGGAGTGTATTTCTATGTGGCAGCTCTGGCATTGACACTAACCTGCTAACT 3055
Qy 971 GlyGlyPheThrTrpLeuCysIleCysCysLysArgGlnLysArgThrLysIleArg 990
Db 3056 GGAGCTGTGCTTGGCTGTGCATCTGTCTGTCAGGAGACGGAAGAGACTAAGTCCGG 3115
Qy 991 LysLysThrLysTyrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArg 1010
Db 3116 AAGAAAAACCAAGTATACCATCTCAGACAGCATGACGACGAAACAGGAAAGAAATGGAACTGAGG 3175
Qy 1011 ProlysTyrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGlu 1030
Db 3176 CCCAAGTACCGGTATCAAGCAGGAGCAGGAGCAGCAACTCCAGTCTGATGGTGTGCGAG 3235
Qy 1031 SerGluPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsn 1050
Db 3236 TCTGAGTTTGAACGACACAGGACACGCTATTACGCCGAGAGAGGATGGAGAGAGGAGTC 3295
Qy 1051 ProlysValSerMetAsnGlySerIleArgAsnGlyAlaSerPheSerTyrCysSerLys 1070
Db 3296 CTGAAGGGTTCCCTGAATGGCTGTGCCAGAAATGGAGTTTCTTCGGTTACTACTCAAAG 3355
Qy 1071 AspArg 1072

Db 3356 GACAGA 3361

RESULT 2
AK084668
LOCUS
DEFINITION
Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched homolog [Homo sapiens], full insert sequence.

ACCESSION
AK084668
VERSION
AK084668.1
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636

REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20493974
PUBMED
11042159

REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861

REFERENCE
4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 695-690 (2001)
PUBMED
11768611

REFERENCE
5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
REFERENCE
6 (bases 1 to 3660)

REFERENCE
7 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hoshizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, D., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, R., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222).

COMMENT
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES
source
1. 3660
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:D330027H07"
/db_xref="taxon:10090"
/clone="D330027H07"
/tissue type="heart"
/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="13 days embryo"
100..3357
/note="unnamed protein product; HYPOTHETICAL 103.9 KDA
PROTEIN homolog [Homo sapiens] (SPTRIAL55781, evidence: PASTY, 85.5%ID, 97.4%length, match=2763)
putative"
/codon_start=1
/protein_id="BAC39244.1"
/db_xref="GI:26351215"
/translation="MEKGLGVKPSASWLVPGYCWTSVKLPRSLYLILYSPFCFSVLW
LSTDADESRQCKTLYAGLRTGEGHLLRLLAGSLPHACRAACCRDSACHALMWLE
GMCFQADCSQSPQFPRTSSNMLII FQKSTQTTDDIGLLPDEDEPHLLRGWRTS
WRQSLGAPLTLSVPSHSHQSLDRQKRLSVVPTHGAMQHSKVNSEAGALSPDA
SAEVRKITVAGSFTSNHTQTPKPKNSVTHRPSHSVPQVKSSTHSPSTDA
PLPVASYSYATPQQASQTSAPHPKELVVSAGKSVQITLPKRVQLNAFLVPE
ABPGSTYDMQLTHPTSDYSGEVHKSQSLKSLTPGLYEFVTVDDQNAHGEGY
VNVTPKPKRPPVAVSPQFQESLFTTSDGATNSTASLTWNKAVDYPVNAVAGP
REERISDITALLKSLVPGNYTSLFTVSDGATNSTASLTWNKAVDYPVNAVAGP
NOVTLTPQNSITLFGNSTDDHGTTSYEWLSPPSKGVEMQGVRTFALOLASMOEG
DYQLVLTDTAGQQAQTVTVIQPENKPPQADAGPKELTLPVDSVTTLDGSKSTD
DQVVSVMNEQSGPDGVLNANSSVATVGLQGVYVTLTVKDERNLQSSQSVNV
IVKEEINKPPVAKIAGNVVTLPTSTAEGLDSRSSDDKGIIVSYLWTRDETSPAGEVL
NHSDDHPVLFLNLVEGTYTHLKVTDKAGSDTDTTVEKPKRNSLVBIIIDVN
EVAAMKSELQKQAKFLIPLALBISTVTCOLGSDHGHCDSTFKRCVCDPFWMENFI
KVQLRDGSDNCSEWLVVLIASFVIVVALGILSWTTCCKKQKQKPKRKYKILDA
TKQSLKLPKTSRAGSKQKPTLSLSLMSHSESLDSDAIFTVPDREKGLLAHWC
SQSLQHGHWCRAPVCHETSTNSTPGSDSGFCRGGKMLSSDYRTIRTKLSALVHGK"

ORIGIN
Alignment Scores:
Pred. No.: 6,48e-207 Length: 3660
Score: 2420.50 Matches: 524
Percent Similarity: 62.16% Conservative: 138
Best Local Similarity: 49.20% Mismatches: 295
Query Match: 43.38% Indels: 109
Gaps: 19
US-10-764-390-3 (1-1072) x AK084668 (1-3660)
QY 10 SerLeuLeuLeuValThrIleAlaGlyCysAlaArgLysGlnCysSerGluGlyArg 29
|||||
Db 220 AGCGTCTGCTGTTGTCACAGATGCTCATGAGAGCAGAGAGAGAGAGAG 273
|||||
QY 30 ThrTyrSerAsnAlaValIleSerProAsnLeuGluThrThrArgIleMetArgValSer 49
|||||
Db 274 ACATTTTATGAGCTGGCTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 330
|||||
QY 50 HisThrPheProValValAspCysThrAlaAlaCysCysAspLeuSerSerCysAspLeu 69
|||||
Db 331 GGAAGCTGCTGCTTCCACAGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCTGGCTGCT 390
|||||
QY 70 AlaThrTyrPheGluGlyArgCysTyrLeuValSerCysProHisLeuGluAsnCysGlu 89
|||||

391	Db	CTATGGTGGCTGGAGGAGTGTGCTTTTCAGGCTGACTGACAGTAAGAGCCCAAGAGAGCTGCCAG	450
90	Qy	ProLysLysMetGlyProIleArgSerSerLeuThrPheValLeuArgProValGlnArg	109
451	Db	CTTTTAGGACAGACTCTTCCAATTCGCTG	483
110	Qy	ProAlaGlnLeuLeuAspTyrGlyAspMetMetLeuAsnArgGlySerProSerGlyIle	129
484	Db	-----ATCATTTTCAAAAATCCCAAACTACAGATGAT	516
130	Qy	TpGlyAspSerProGluLysAspIleArgLysAspLeu**PheLeuGlyLysAspTrpGly	149
517	Db	TTGGGCTTCTGCTGAAGATGATGAACCAATCTTCTGAGGCTAGGC-----TGGGCC	570
150	Qy	LeuGluGluMetSerGluTyr***AspAspTyrArgGluLeuGluLysAspLeuLeuGln	169
571	Db	AG-GACATCTGGAGGAG-----GCAGAG	593
170	Qy	ProSer---GlyLysGlnGluPro-----Arg	177
594	Db	CTTCTTTGGGGCTCCCTCCACCTTTCTGTACCCCTTAGTCACCCAGAGCTTACTCAG	653
178	Qy	GlySerAla---GluTyrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPhe---	195
654	Db	GGATCGCAGAGAGAGATCTCAGTGTGGTACCTACATGAGCGGAGCAGCATCTCAA	713
196	Qy	AsnSerSerValGlyAspSerProAlaValProAlaGluThrGlnGlnAppProGluLeu	215
714	Db	AGTGAATCACTCGGAGGAGCGTGTCTCAGTCCCACTCTGCAGAGGTCCGCAAAAC	773
216	Qy	HisTyr-----LeuAsnGlu-SerAlaSerThrProAlaProLysLeuProGln	231
774	Db	CATTACAGTTGCCGCTTCCTTCCACGATAACCACTACACAGACTCTCAGTGGCCCAA	833
231	Qy	uArgSerValLeuLeuProLeuProThrThrProSerSerGlyGluValLeuGluLysGln	251
834	Db	GAATGTCTCCATCCATCCTGAACATCC-----	861
251	Qy	uLysAlaSerGlnLeuGlnSerSerAsnSerSerGlyLysGluValLeuMetPr	271
862	Db	-----GAGCACTCCAGTCTCTGTATCTGT-----	885
271	Qy	oSerHisSerLeuProProLaserLeuGluLeuSerSerValThrValGluLysSerPr	291
885	Db	-----	885
291	Qy	oValLeuThrValThrProGlySerThrGluHisSerIleProThrProProThrSerAl	311
886	Db	-----ACTCCGCAAGTAAAGACACTGAGCAGACACTGATGCCCTCTGCCAGT	938
311	Qy	aAlaProSerGluSer-----ThrProSer---GluLeuProIleSerProThrThrAl	328
939	Db	GGCCCCCTCTACAGCTATGCCACCCCGACCCCGAGCCCTTCTCAGAGCACTCAGC	998
328	Qy	aProArg---ThrValLysGluLeuThrValSerAlaGlyAspAsnLeuIleIleThrLe	347
999	Db	ACCACACCCAGTTGTAAAGGAGCTGGTGTCTGCTGCGAAGAGCGTCCAGATCACCT	1058
347	Qy	uProAspAsnGluValGluLeuLysAlaPheValAlaProAlaProProValGluThrTh	367
1059	Db	GCCTAAGAATGAAGTTCAGTTAAATGCCTCGTCTTCCGAAGCAGAGCCAGGAGAAAC	1118
367	Qy	rTyrAsnTyrGluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGluIleLysGln	387
1119	Db	CTACACCTACAGCTGGCAGCTGATCACTCATCTTACAGACTACAGTGGAGAGGTGGAGAG	1178
387	Qy	nGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysVa	407
1179	Db	GAACATCTCCAGAGCTCCACATGTCTCCAGCTGACTCCAGGCTGTACGNAATTCAGGT	1238
407	Qy	lThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAl	427
1239	Db	GACTGTGGATGGCCGAATGCCATGGGAGGCTATGTGAATGTGACAGTGAACACAGA	1298

QY 787 rPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThrValG1 807
 Db 2379 GTTTCACCTGAAAGTGACAGATGCAAGGCGGAGCGACACAGACGACAGTGGGA 2438
 QY 807 uValGlnProAspArgLysSerGlyLeuValGluLeuThrLeuGlnValGlyValG1 827
 Db 2439 AGTGAAGCCTGACCCCGAGAAAGCAACCTAGTGGAGATCATCTTGGATGTGAACGTGAC 2498
 QY 827 yGlnLeuThrGluGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuAsnVa 847
 Db 2499 TCAGCTGACTGAGAGCTGAAGGAGATGCTCATCCGCGAGATTTGGGTCTCTCTGGGGT 2558
 QY 847 lleuAspSerAspIleLysValGlnLysIleArgAlaHisSerAspLeuSerThrValI1 867
 Db 2559 GCTGGATCCGACATCATTTGTCMAAAGATTACCGCGTACAGCGAGACAGACCAAGAT 2618
 QY 867 eValPheThrValGlnSerArgProPheLysValLeuLysAlaGluValAlaAr 887
 Db 2619 GTTGTGTTTGTTCAGAACGACCTCCACCAGCTCTTCAAGAGCCATGAGGTGGCAGC 2678
 QY 887 gAsnLeuHisMetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuAr 907
 Db 2679 CATGCTCAAGACGAGCTGCAGAGCAGAGGCTGACTTCTCATCTTCAGAGCCCTGGA 2738
 QY 907 gValAspThrAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuTh 927
 Db 2739 ATTCAGCAGCTGACATGTCAGCTGACCTGTTCTGACCATGGCCACTGTGACTCATTCAC 2798
 QY 927 rLysArgCysIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyrlTrpAs 947
 Db 2799 CAAGCGCTGTGTCTGACCCGTTTTGGATGGAGAATTTTCATCAAGGTGCAGCTGAGGGA 2858
 QY 947 pGlyLeuSerAsnCysGluTrpSerIlePheTyrlValThrValLeuAlaPheThrLeuI1 967
 Db 2859 TGGAGCAGCACTGTGAATGAGCGTGTCTACGTCATCATCTGCTCTTTGTCTATGT 2918
 QY 967 eValLeuThrGlyGlyPheThrTrpLeuCysIleCysCysLeuArgGlnLysArgTh 987
 Db 2919 TGTTCCTTGGGATCCTGTCATGACTACATCTGCTGCTGCAAGAGCCAAAAA--GG 2975
 QY 987 rLysIleArgLysLysThrLysTyrlThrIleLeuAspAsnMetAspGluGlnGluArgMe 1007
 Db 2976 AAAACCCAGAGAAAGCAGATACCAAGATCTCTGGATGCCAGAT---CAGGAGAGCCT 3032
 QY 1007 tGluLeuArgPro-----LysTyrlLysLysHisArgSerThrGluHisAsnSerSe 1025
 Db 3033 GGAGCTGAAACCAACCTCCGAGCAGGAGCAAGAAAGCCCAAGCCAGCTGAGCAGCAG 3092
 QY 1025 rLeuMetValSerGluSerGluPheAspSerGlnAspThrIlePheSerArgGluLyl 1045
 Db 3093 CTGTATGCTATTTGATCGAGCTGACAGCGAC---GATGCCATCTTCATATGCCCGAGA 3149
 QY 1045 sMetGluArgGly 1049
 Db 3150 CCGGAGAGAGGC 3162
 RESULT 3
 LOCUS AK043006 4141 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:A730047D20 product:HYPOTHETICAL 103.9 KDA PROTEIN homolog (Homo sapiens), full insert sequence.
 ACCESSION AK043006
 VERSION AK043006.1 GI:26335462
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)

99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
 6 (bases 1 to 4141)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.
 Location/Qualifiers
 1. 4141
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:A730047D20"
 /db_xref="taxon:10090"
 /clone="A730047D20"
 /tissue_type="cerebellum"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"

FEATURES
 source

```

/dev_stage="7 days neonate"
205_..3462
/note="unnamed protein product; HYPOTHETICAL 103.9 KDA
PROTEIN homolog [Homo sapiens] (SPTAAL55781, evidence:
FASTY, 85.5%ID, 97.4%length, match=2763)
putative"
/codon_start=1
/protein_id="BAC31432.1"
/db_xref="GI:26339463"
/translation="MEKRLGVKPSASVWLPGYCMQTSVKLPSRLYLILYSPFCSVLW
LSTDADSRCCQGGKTLVAGLRTEGNEHLRLAGSLPFAHCAACCRDSDACHALWLM
GMCQADSCSKQSPPTDSSNMLIIFOKQSTDDGLGLPDEDEPHILLRGLWRTS
WRROSLGAPLTLVPSSSHQSLDRDOKRDLVVPVTHGAMQHSKVNSHREAGALSPT
SAEVRKTIIVAGSTSHHTTQTTPWPKNVSIHPSESHSPVSGTPQVKSTESPTDA
PLVPASPTIATPQASQTSAPHPVVKSLVSGAKSVQITLIPKNEQLNAPLPE
AEPGETTYDQWLTITPDYSGEVRKHSQSLKLTPLGLYEFKVTVDGQNAHGGY
LVNVTPEPRKRPVAVVSPQFOEISLPTTSTIIDGQSDDDDKIYQVHWEELKGPL
REEKISDATALKLSKLVPGNYTSLTVDSGDGATNSTASLTWNKAVDYVPVNAOG
NOVITLPONSITLFGNSTDGHGITSVEMSLSPSSKGVNEMQGVRTPALOLSAMOG
DYTQLTVDTAGQAQNAVTVIQQPENKPPQADAGDPDKELTLPVSTLIDGKSTD
DQRVSYLWESQSDGQVLENANSSVAITVGLQGVITVFTLVKDRNLQSSVNV
IVRKEINPKPVAKIAGNVVVTLPSTAEILDGSRSDDKGIVSYLWTRDNTSPAAEVL
NHSDEHPLVFLSNLVEGTYTFLKVDIAKESDTRTVEKVPDPKSNLVEIILDVN
VSQTLERLKLMIROIGVLGLVDSDIIVQIKPTQEOSTKMLFVQNDPPHOLPKGH
EVAAMLKSELQKADFLIIRALRISTVTCLNCSDHGCHDSFTKRCVDCPFWMENFI
KQVLRDGSDNCEWSLVYIIASPVIVVALGILSWTITCCCKRQKGRKPRKRYKILDA
TDQSELEUKPTSRAGSKOKGPTLSSLMHSESLSDDAIFWDPREKGLULALHWCR
SQSLHHTGWCRAVPVCHETSTNSTPGSDSGFCRGGKRLSSDYRTRTKSALVHGK"
4119...4124
polyA_signal
/note="putative"
4141
polyA_site
/note="putative"
4141
ORIGIN
Alignment Scores:
Pred. No.: 7,94e-207 Length: 4141
Score: 2420.50 Matches: 524
Percent Similarity: 62.16% Conservative: 138
Best Local Similarity: 49.20% Mismatches: 295
Query Match: 43.38% Indels: 109
DB: 3 Gaps: 19
US-10-764-390-3 (1-1072) x AK043006 (1-4141)
QY 10 SerLeuLeuLeuValThrIleAlaGlyCysAlaArgLysGlnCysSerGluGlyArg 29
Db 325 AGCGTTCTGTGTTCTCAACAGATGCTGATGAGACAGA-----TGCCAACAGGGGAAG 378
QY 30 ThrTyrSerAsnAlaValIleSerProAsnLeuGluThrThrArgIleMetArgValSer 49
Db 379 ACACTTTATGAGCTGGCTTGAGAACTCGAGGAGAA---AATCACTCCCGCTCTCTTGA 435
QY 50 HisThrPheProValValAspCysThrAlaAlaCysCysAspLeuSerSerCysAspLeu 69
Db 436 GGAACCTGCTTTCCACGCTTTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495
QY 70 AlaTrpTrpPheGluGlyArgCysTyrLeuValSerCysProHisLysGluAsnCysGlu 89
Db 496 CTATGCTGGCTGGAAGGATGCTTTCAGGCTGACTGCAGTAAGCCCCAGAGCTGCCAG 555
QY 90 ProLysLysMetGlyProIleArgSerTyrLeuThrPheValLeuArgProValGlnArg 109
Db 556 CTTTATAGGACAGACTCTTCCAAATTCATGCTG----- 588
QY 110 ProAlaGlnLeuAspTyrGlyAspMetMetLeuAsnArgGlySerProSerGlyIle 129
Db 589 -----ATCATTTTCAAAATCCCAACTACAGATGAT 621
QY 130 TrpGlyAspSerProGluAspIleArgLysAspLeu***PheLeuGlyLysAspTrpGly 149
Db 622 TTGGGCTTCTGCTGAGATGATGAACACATCTTCTGAGGCTAGGC-----TGGGGC 675
QY 150 LeuGluGluMetSerGluTyr***AspAspTyrArgGluLeuGluLysAspLeuGln 169

```

Db 1584 CATACTAAACTTAGTAAGTCTGTCCTCCCGGGGAACCTACACCTTTCAGCTTAACCTGTTGTGCGA 1643

Qy 507 pSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuLeuValAsnAsnAlaValAspTy 527

Db 1644 CTCTGACGGGGTACCAACTCCACACTGCAAGCTGACTGTGAACAAGCTGTGGACTA 1703

Qy 527 rProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIleTh 547

Db 1704 CCTCCCGTGGCAATGTCAGGCCCCAACCAAGTGTATCACCTGCTCGAAGCTCCATCAC 1763

Qy 547 rLeuAsnGlnAsnGlnSerSerAspAspHisGlnIleValLeuTyThrGluTrpSerLeuGl 567

Db 1764 CCTCTTTGGAAACCGAGCAGCATGACCAACCGGCATCACAGCTATGAGTGTGCTCGAC 1823

Qy 567 yProGlySerGluGlyLysHisValValMetGlnGlyValGlnThrProTyLeuHisLe 587

Db 1824 CCGGAGCAGCAAGGAAGGTGGTGGAGATGACGAGGAGTTGAACGCCAGCCCTCGAGCT 1883

Qy 587 uSerAlaMetGlnGluGlyAspTyThrPheGlnLeuLysValThrAspSerSerArgGl 607

Db 1884 GTCCGCAATGCAAGGAAGGAGACTATACCTACCAGCTTCACAGTGTACACCGCGAGGACA 1943

Qy 607 ngInSerThrAla***ValThrValIleValGlnProGluAsnAsnArgProProValAl 627

Db 1944 ACAGGCCACCGGCCAAGTGAAGTGTGATTTGTGAGCTGTGAGAACAAACAGCCCTCTCAGGC 2003

Qy 627 aValAlaGlyProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGlySe 647

Db 2004 AGATGAGCCCGCCAGCAAGAGCTGACCTGCCGCTGGGAGACGACACACCTCGAGCGGAG 2063

Qy 647 rSerSerSerAspAspHisGlyIleValPheTyThrHisTrpGluHisValArgGlyProSe 667

Db 2064 CAAGAGCACAGATGACACAGAGTGTCTCTTACCTTTGGGAGCAGAGTGGGGACCTGA 2123

Qy 667 rAlaValGluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnValGl 687

Db 2124 CGGGGTGTCAGCTGGAGATGCCAACAGACAGTGTGCGCACTGTGAGTGGGCTGCAAGTGG 2183

Qy 687 yThrTyHisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrSerThrLe 707

Db 2184 GACTTATGTATTACCTTGTACCTGTCAAGATGAGAGGAACCTTACAGAGCAGAGTCTCGT 2243

Qy 707 uThrValAlaValLysGluAsnAsnSerProProArgAlaArgAlaGlyGlyArgHi 727

Db 2244 TAATGTCTATTGTCAAGAGAAATAAACAAACCGCCAGTAGCAAGATCGCTGGGAAGCT 2303

Qy 727 sValLeuValLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAspAspGl 747

Db 2304 GGTGTGTGACCTTGCCCAACGAGCACAGCTGAGTGTGAGTGTGAGTGTCTCAGATGACAA 2363

Qy 747 nArgIleValSerTyLeuTrpIleArgAspGlyGlnSerProAlaAlaGlyAspValIl 767

Db 2364 GGGGATGATGACTACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2423

Qy 767 eAspGlySerAspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyValTyThr 787

Db 2424 GAATCACTCTGACCAACCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2483

Qy 787 rPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThrValGl 807

Db 2484 GTTTCACCTGAAAGTACACAGATCAAAAGGGCGAGAGCGACACAGACCGGACGACAGTGA 2543

Qy 807 uValGlnProAspProArgLysSerGlyLeuValGluLeuThrLeuGlnValGlyValGl 827

Db 2544 AGTGAAGCTTGACCCCGAGGAAAGCAACCTTAGTGGAGATCACTTGGATGTGAACGTGAC 2603

Qy 827 yGlnLeuThrGluGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnVa 847

Db 2604 TCAGCTGACTGAGAGGCTGAAGGGATGCTCATCCCGCAGATTTGGGGTCTCTCTGGGGT 2663

Qy 847 lLeuAspSerAspIleLysValGlnLysIleArgAlaHisSerAspLeuSerThrValIl 867

Db 2664 GCTGGATTCCGACATCATTTGTGAAAAAGATTTCAGCCGTATACCGGACGAGCAGCACAAGAT 2723

Qy 867 eValPheTyValGlnSerArgProProPheLysValLeuLysAlaAlaGluValAlaAr 887

Db 2724 GTTGTGTTTGTTCAGAACGACCTCCCCACAGCTCTTCAAGGCCATGAGTGGCAGC 2783

Qy 887 gAsnLeuHisMetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuAr 907

Db 2784 CATGCTCAAGAGCGAGCTGCAGAGACGAGCTGACTTCTCTCATCTTTCAGAGCCCTGGA 2843

Qy 907 gValAspThrAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuTh 927

Db 2844 AATCAGCACATCATGTGAGCTGAACCTGTTCAGCATCGCCACTGTGACTTCAATTCAC 2903

Qy 927 rLysArgCysIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyIleTrpAs 947

Db 2904 CAAGCGCTGTGTGTGACCGCTTTTGGATGGAGAAATTCATCAAGGTGCAGCTGAGGA 2963

Qy 947 pGlyCysSerAsnCysGluTrpSerIlePheTyThrValThrValLeuAlaPheThrLeuIl 967

Db 2964 TGGAGACAGCAACTGTGAATGGAGCGTCTCTACGTCTCATCTTGTCTTGTCTATTGT 3023

Qy 967 eValLeuThrGlyGlyPheThrTrpLeuCysIleCysCysLysAspGlnLysArgTh 987

Db 3024 TGTTCCTTGGGATCCTGTCTGATGACTACATCTCTGCTGCAAGAGCGCAAAA---GG 3080

Qy 987 rLysIleArgLysLysThrLysTyThrIleLeuAspAsnMetAspGluGlnLysArgMe 1007

Db 3081 AAAACCCCAAGAGAAAGACAGATACAGATCTTGGATGCCACAGAT---CAGGAGAGCCT 3137

Qy 1007 tGluLeuArgPro-----LysTyGlyIleLysHisArgSerThrGluHisAsnSerSe 1025

Db 3138 GGAGCTGAACCAACCTCCCGAGCAGGAGCAACACAGAAAGGCCCCACGCTGAGCAGCAG 3197

Qy 1025 rLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIlePheSerArgGluLy 1045

Db 3198 CCTGATGCATTCTGAATCGAGCTGGACAGCGAC---GATGCCATCTTCACATGGCCAGA 3254

Qy 1045 sMetGluArgGly 1049

Db 3255 CCGGAGAGAGGC 3267

RESULT 4

AF289597 3616 bp mRNA linear HTC 01-JAN-2002

LOCUS Homo sapiens clone pp791 unknown mRNA.

DEFINITION AF289597

ACCESSION AF289597

VERSION AF289597.1 GI:18027397

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Qin, W.X., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Huang, Y., Zhao, X.T., Wan, D.P. and Gu, J.R.

TITLE Novel human cDNA clones with function of inhibiting cancer cell growth

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3616)

AUTHORS Qin, W.X., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Huang, Y., Zhao, X.T., Wan, D.P. and Gu, J.R.

TITLE Direct Submission

JOURNAL Submitted (17-JUL-2000) National Laboratory For Oncogenes & Related Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai 200032, P. R. China

FEATURES

Location/Qualifiers

source 1..3616

organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="pp791"

133..2973

/codon_start=1

CDS

```

/product="unknown"
/db_xref="GI:18027398"
/translation="MCQADCSRPSQCFRAFTHSSNMLVFLKKFQTADDLGLFLPEDD
VPHLLGLGNWASNRQSPRAALRPVSSDQSLIRLQKRGSFSDVVTIVTQHSK
VNDNELGLTTSGSAEVRHKAITISSPLTDLTAELSGGPKNSVQPEISRLGATPS
TQOVSKSEKTOIAVQPVAPYSVATPTQASFOSTSAFYVPIKELVVSAGESVQITL
PKNEVLNAYVLQPPKGETTYDNQLITHPRDYSGMEGKHSQILKSLKTPGLYEP
KIVQVQNAHGEGVNVTKPEPRKNRPPIAIVSPQFQELISLPTISTVIDSGSTDD
KIVGHWEELKGLPREEKISDITAILKSLKLPNGNYFSLTVDSGDATNSTANLTV
NKAVDYPVANAGNQVITLPQNSITTLFNGNSTDDHGITSYEWLSPSKGVEMQOG
VTRTLOLSAMOEGDYTVLTQVTTDTIGQOATQVTVIROPENKPPQADAGPKDELTL
PVDSTLQSGSKSDDKITISVLEWKTQGPDCGVOLNANSSVATVGLQGVTVFTLV
KDERNLQSQSVNVIKEEINKPPIAKITGNVITLPRSTAEILDGSKSDKGLVSYL
WDRGSPAAEGEVNHSDDHPIFLSNLVEGTYTFHLKVTDAKESDSDRTITVBKDP
PRKNLVEIILIDINVQSTLERLKMFRQIGVLGLVLDSDIIVQIKIPTYEQSTKMF
FVQNEPHQIFKGEHVAAMLKSELKQKADLELI FRALENTVITCOLKSDHGHCDSFT
KRCTCDPWFENFTKVQLRDGDNSCNESLVYIATFVIVVALGILSWTVICCKKROK
GPKRKSKYKTLDTADQESLSLKPTSRAGIKQKGLLLSSLMHSESELDSDDAITFWP
DREKGLLHGQGVSPNGQTPFKARSPEEIL"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 4,04e-204 Length: 3616
Score: 2389.50 Matches: 509
Percent Similarity: 61.26% Conservative: 125
Best Local Similarity: 49.18% Mismatches: 288
Query Match: 42.82% Indels: 113
DB: 3 Gaps: 15

```

US-10-764-390-3 (1-1072) x AF289597 (1-3616)

```

QY 57 CysThrAlaAlaCysCysAspLeuSerSerCysAspLeuAlaTrpTrpPheGluGlyArg 76
DB TGTGGGCTGCTGCTGCAGACTCTGCCTGCCATGCTCTTTTGGTGCTAGAAGGAGT 135
QY 77 CysTyrLeuValSerCysProHisLysGluAsnCysGluProLysLysMetGlyProle 96
DB TGCATTCAGGCAGACTGCAGCAGGCCCCAGAGCTGCCGGCTTTTAGGACACACTCTCC 195
QY 97 ArgSerTyrLeuThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyr 116
DB AATTCATCTGCTGCTTTT---TTAAAAAATTCCAACTGCAGATGATTG----- 243
QY 117 GlyAspMetMetLeuAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAsp 136
DB -----GGCTTTCTACTGAAGAT 261
QY 137 IleArgLysAspLeu***PheLeuGlyLysAspTrpGlyLeuGluGluMetSerGluTyr 156
DB GATGTACACATCTCTCGGGCTAGGTTGGAGCTGGCA----- 300
QY 157 **AspAspTyrArgGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluPro 176
DB -----TCTTGAGGAGCAGGCCA 318
QY 177 ArgGlySerAlaGluTyrThrAspTrpGlyLeuLeuProGlySerGluYalaPheAsn 196
DB ----- 333
QY 197 SerSerValGlyAspSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHis 216
DB -----AGCTTAATCAGGAAGCTTCAGAGAGAGGTTAGTCCCAAGT 402
QY 217 TyrLeuAsnGluSerAlaSerThrProAlaProLysLeuProGluArg----- 232
DB ----- 364
QY 233 SerValLeuLeuProLeuProThrProSerSerGlyGluValLeuGluLysGluLys 252
DB -----AAAGTGAATGACTCCACGAA 456
QY 253 AlaSerGlnLeuGlnGlnSerSerAsnSerSerGlyLysGluValLeuMetProSer 272

```

```

DB 457 TTAGGTGCTGTGACTACACAGTGGCTCTGCAGAGGTCCACAGGCGATTACAAITTCAGT 516
QY 273 HisSerLeuProProAlaSerLeuLeuLeuSer-----SerValThr 286
DB 517 CCCCTAAACACAGACCTGACTGCAGAGCTGTCTGGTGGGCCAAAAGAAATGTATCAGTCAA 576
QY 287 ValGluLysSerProValLeuThrValThrProGlySer----- 299
DB 577 CCTCAAAATATCAGAGGTCTTGTCTACTACGCCCCAGCACCAACAGTAATAAAGTTCTGAG 636
QY 300 ---ThrGluHisSerIleProThrPro-----ProThr 309
DB 637 AAAACCCAGATTGCTGTGCTCCACCCAGCAGTGGCTCCCTCTACAGTTATGTCTACCCCTACC 696
QY 310 SerAlaAlaProSerGluSerThrProSerGluLeuLeuProleSerProThrThrAlaPro 329
DB 697 CCCAGGCGCTCTTCCAGAGCAGCCTCAGCACCATACCCAGTT----- 738
QY 330 ArgThrValLysGluLeuThrValSerAlaGlyAspAsnLeuIleLeuThrLeuProAsp 349
DB 739 -----ATAAAGGAACCTGGTGTATCTGCTGGAGAGAGTGTCCAGATAACCTGCCTAAG 792
QY 350 AsnGluValGluLeuLysAlaPheValAlaProAlaProValGluThrThrTyrAsn 369
DB 793 AATCAAGTTCAATTAATGATATGTTCTCAAGAACCACTTAAAGGAGAACCTACACC 852
QY 370 TyrGluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGluIleLysGlnGlyHis 389
DB 853 TACGACTGGCAGCTGATTACTCATCTAGAGACTACAGTGGAGAAATGGAAGGGAACAT 912
QY 390 LysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrVal 409
DB 913 TCCAGATCTCTCAAACTATCGAAGCTCACTCCAGCCTGTATGAATTCAAAGTGAATTGA 972
QY 410 SerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysProAlaArgArg 429
DB 973 GAGGCTCAAAATGCTCCATGGGAGGCTATGTGAACGTGACAGTCAGCCAGAGCCCGT 1032
QY 430 ValAsnLeuProProValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeu 449
DB 1033 AAGAATCGGCCCCCATTTGCTATTGTGTCACTCAGTTCCAGGAGATCTCTTTGCCAAC 1092
QY 450 ThrSerAlaLeuIleAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHis 469
DB 1093 ACTTCTCAGTCATTGATGGCAGTCAAAGCACTGATGATGATAAATCGTTTCAGTACCAT 1152
QY 470 TrpGluGluIleAsnGlyProPheIleGluLysThrSerValAspSerProValLeu 489
DB 1153 TGGGAAGAACCTTAAGGGCCCTCTAAGGAAGAGAGATTCTCTGAAGATACAGCCATATTA 1212
QY 490 ArgLeuSerAsnLeuAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAsp 509
DB 1213 AAACCTAAGTAACCTCGTCCCTGGGAACCTACACTTTTCAGCTTGTAGTAGACTCTCAT 1272
QY 510 GlyAlaThrAsnSerThrThrAlaAlaLeuIleValAsnAlaValAspTyrProPro 529
DB 1273 GGAGCTACCAACTCTACTACTGCAACCTGACAGTGAACAAAGCTGTGGATTACCCCCCT 1332
QY 530 ValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsn 549
DB 1333 GTGGCCAAACGAGGCCCCCAACAAAGTATACCTGTGCCCAAACTCATCACCCCTCTTT 1392
QY 550 GlyAsnGlnSerSerAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGly 569
DB 1393 GGGAAACGAGACCTGATGATCATGGCATCACAGCTATGATGATGATGATGATGATGATGAT 1452
QY 570 SerGluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAla 589
DB 1453 AGCAAAAGGAAAGTGTGGAGATCAGGCTGTAGAACACCAACCTTACAGCTCTCTCGG 1512
QY 590 MetGlnGluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSer 609

```


[illegible]

them based on alignment.

```
FEATURES              Location/Qualifiers
   source             1..2959
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10090"
gene                  <1..>2959
                        /locus_tag="HCM6309"

ORIGIN
Alignment Scores:      2.06e-167          Length:           2959
Pred. No.:            1982.00           Matches:         448
Score:                55.32%            Conservative:    119
Percent Similarity:   43.71%            Mismatches:     353
Best Local Similarity: 35.52%            Indels:        106
Query Match:         9                 Gaps:          17

US-10-764-390-3 (1-1072) x AY417695 (1-2959)

Qy      10 SerLeuLeuLeuValThrIleAlaGlyCysAlaArgLysGlnCysSerGluGlyArg 29
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      121 AGCGTTCTGTGTGTCAACAGATGCTCATGAGACAGA-----TGCACACAGGGGAAG 174

Qy      30 ThrTyrSerAsnAlaValIleSerProAsnLeuGluThrThrArgIleMetArgValSer 49
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      175 ACATTATTATGGAGCTGGCTTGAGAACTGAGGGAGAA---AATCACCTCCGGCTTCTTGCA 231

Qy      50 HisThrPheProValValAspCysThrAlaAlaCysAspLeuSerSerCysAspLeu 69
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      232 GGAAGCCTGCCCTTTCCACGCGCTGCGGGCTGCCCTGCCGGACTCTGCGCTGCCACGCT 291

Qy      70 AlaTrpTrpPheGluGlyArgCysTyrLeuValSerCysProHisLysGluAsnCysGlu 89
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      292 CTATGGTGGCTGGAAGGATGTGTTTTTCAGCGCTGACTGCAGTAAGCCCCAGAGCTGCCAG 351

Qy      90 ProLysLysMetGlyProIleArgSerTyrLeuThrPheValLeuArgProValGlnArg 109
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      352 CCTTTTAGGACAGACTCTTCCAATTCCATGCTG-----384

Qy      110 ProAlaGlnLeuLeuAspTyrGlyAspMetMetLeuAsnArgGlySerProSerGlyIle 129
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      385 -----ATCAATTTTTCAAAATCCAAACTACAGATGAT 417

Qy      130 TrpGlyAspSerProGluAspIleArgLysAspLeu***PheLeuGlyLysAspTrpGly 149
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      418 TTGGGCCTTCTGCTTGAGATGATGAACCACCATCTTCTGAGGCTAGGC-----TGGGGC 471

Qy      150 LeuGluLeuMetSerGluTyrr**AspAspTyrArgGluLeuGluLysAspLeuLeuGln 169
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      472 AG-GACATCTGGGAGGAG-----GCAGAG 494

Qy      170 ProSer---GlyLysGlnGluPro-----Arg 177
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      495 CTTTCTTTGGGGCTCCCTCCACCCCTTCTGTATCCCTCTAGTCACCACAGAGCTTACTCAG 554

Qy      178 GlySerAla---GluTyrrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPhe--- 195
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      555 GGATCGGCAGAAGAGAGATCTCAGTGTGGTACCTACATCATGGAGGCGATGCAGCATTCATA 614

Qy      196 AsnSerSerValGlyAspSerProAlaValProAlaGluThrGlnGlnAspProGluLeu 215
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      615 AGTGAATCACTCGGAGGAAGCAGGTGCTCTGATGCCACCTCTGCAGAGGTCGCGCAAAAC 674

Qy      216 HistYr-----LeuAsnGlu-SerAlaSerThrProAlaProLysLeuProGlu 231
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      675 CATTACAGTTCCGGTTCCTTCCACGAGTAACCACTACACAGACTCTCTGAGTGGCCCAA 734

Qy      231 uArgSerValLeuLeuProLeuProThrThrProSerSerGlyGluValLeuGluLysGlu 251
       |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      735 GAATGTGTCCATTCATCTCTGAACCATTC-----762

Qy      251 uLysAlaSerGlnLeuGlnGlnSerSerAsnSerSerGlyLysGluValLeuMetPr 271
```



```
Qy 414 aPheGlyGluGlyPheValAsnValThrValIlyProAlaArgValAsnLeuProPr 434
Db 1164 CCATGGGGAAGGCTATGTGAACGTCGACAGTCAAGCCAGAGCCCGCTAAGAATCGGCCCC 1223
Qy 434 oValAlaValValSerProGlnLeuGlnGluLeuThrLeuThrSerAlaLeu1l 454
Db 1224 CATGTCTATTGTGTCACTCAGTTCAGGAGATCTCTTTGCCAACCACTTCTCAGTCAT 1283
Qy 454 eAspGlySerGlnSerThrAspThrGluLeuValSerThrHisTrpGluGluLeuAs 474
Db 1284 TGATGGCAGTCAAGCAGCTGATGATGATAAATCGTTCAAGTACCATTGGGAAGAACTTAA 1343
Qy 474 nGlyProPheLeuGlnGluLeuThrSerValAspSerProValLeuArgLeuSerAsnLe 494
Db 1344 GGGCCCTCTAAGAGAAGAGAAGATTCTGAAGATACAGCCATATTAAACCTAAGTAAACT 1403
Qy 494 uAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSe 514
Db 1404 CGTCCCTGGGAACCTACACTTTTCAGCTTGATGTAGTACTCTGATGAGCTACCAACTC 1463
Qy 514 rThrThrAlaAlaLeuLeuValAsnAsnAlaValAspTyrProProValAlaAsnAlaGl 534
Db 1464 TACTACTGCAAACTGACAGTGAACAAAGCTGTGGATTACCCCTGTGGCCAAAGCAGG 1523
Qy 534 yProAsnHisThrLeuThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSe 554
Db 1524 CCCCAACCAAGTGTACCTCGCCGCCAAACTCCATCCCTCTTTGGGAACCAAGAGCAC 1583
Qy 554 rAspAspHisGlnIleValLeuThrGluTrpSerLeuGlyProGlySerGluGlyLysH 574
Db 1584 TGATGATCATGGCATCACAGCTAGTGTGTCTACTCAGCCAAAGCAAGGGAAGT 1643
Qy 574 sValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAs 594
Db 1644 GTTGAGATGAGGGGTGTAGAACCAACCTTACAGCTCTCTCGATGCAAGAGGAGA 1703
Qy 594 pTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla***ValTh 614
Db 1704 CTACACTTACCAGCTCAGAGTGTGACACATAGGACAGCAGGCGCACTGCTCAAGTGAC 1763
Qy 614 rValIleValGlnProGluAsnAsnArgProProValAlaValAlaGlyProAspLysGl 634
Db 1764 TGTATTGTGCAACTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1823
Qy 634 uLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerAspAspHisGl 654
Db 1824 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1883
Qy 654 yIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValGluMetGluAsnIl 674
Db 1884 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1943
Qy 674 eAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuTh 694
Db 1944 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2003
Qy 694 rValLysAspGlnGlnGlyLeuSerSerThrThrLeuThrValAlaValLysLysGl 714
Db 2004 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2063
Qy 714 uAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAs 734
Db 2064 AATAACAAACCACTATAGCAAGATAACTGGGAATGTGTGTATCCCTACCCACAGAG 2123
Qy 734 nSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrLeuTr 754
Db 2124 CACAGCAGAGTGTGTGTCTAAGTCTCTAAGTACAGATGACAAAGGGAATAGTACAGTCTG 2183
Qy 754 pIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerVa 774
Db 2184 GACTCGAGATGAGGGGAGCCCGACGAGCGAGGGGAGGTGTAAATCACTCTGACCATCACCC 2243
```

```
Qy 774 lAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThrAs 794
Db 2244 TATCTTTTCTTTTCAAAACCTGTGTGAGGGAACCTACACTTTTCACTTGAAGTGACCGA 2303
Qy 794 pSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArgL 814
Db 2304 TGCAAAAGGCTGAGAGTGCACAGACCGGACCACTGTGGAGGTGAACCTGATCCAGAA 2363
Qy 814 sSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgL 834
Db 2364 AAACAACCTGTGTGAGATCATCTTGGATATCAACGTCAGTCAGCTAAGAGAGGCTGAA 2423
Qy 834 sAspThrLeuValArgGlnLeuAlaValLeuAsnValLeuAspSerAspLysVa 854
Db 2424 GGGGATGTTCAATCCGAGATTGGGCTCTCTGGGGTGTGTGATTCGACATCATTTGT 2483
Qy 854 lGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTyrValGlnSerAr 874
Db 2484 GCAAAAGATTTCAGCCGTACAGGAGCAGACCAAAATGATATTTTGTTCAAACGA 2543
Qy 874 gProProPheLysValLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSe 894
Db 2544 GCCTCCCCACCAGATCTTCAAGGCCCATGAGTGGCAGCGATGCTCAAGAGTGAGCTGCG 2603
Qy 894 rLysGluLysAlaAspPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLe 914
Db 2604 GAAGCAAAAGCGACACTTTTGTATTTATTCAGAGCCCTTGGAGTCAACACTGTGCATGTCA 2663
Qy 914 uLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCysIleCysSerHi 934
Db 2664 GCTGACTGTTCCGACCATGGCCACTGTGATCTGTTTACCAACGCTGTATCTGTGACCC 2723
Qy 934 sLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyGluSerAsnGlyTr 954
Db 2724 TTTTGTGATGAGAAATTCATCAAGGTGCAGCTGAGGATGGAGACAGCAACTGTGAGTG 2783
Qy 954 pSerIlePheTyrValThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheTh 974
Db 2784 GAGCGTGTATATGTTATCATTTGTACCTTTTGTCTTTTGTGCTTGGGAATCTGTCTC 2843
Qy 974 rTrpLeuCysLysCysCysValArgGlnLysArgThrLysIleArgLysLysThrLy 994
Db 2844 TTGACTGTGATCTGTTGTGTAGAGGCAAAA---GGAAACCCAGAGAGAAAGCA 2900
Qy 994 sTyrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro 1011
Db 2901 GTACAAGATCTGTGATCCACGGAT---CAGAAAGCTGGAGCTGAAGCCA 2949

RESULT 9
AY417694
LOCUS
DEFINITION Pan troglodytes HCM6309 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY417694
VERSION AY417694.1 GI:39773654
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 2962)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2962)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
```



```
Db 1517 CATACTAAACTTAGTAACTGCTGCGGGAAGTACACCTTACAGCTTAAGCTGTGTGCGA 1576
Qy 507 pSerAspGlyAlaThrAsnSerThrAlaAlaLeuLeuValAsnAsnAlaValAspTy 527
Db 1577 CTCTGACGGGGCTACCAACTCCACCACTGCAAGCTGACTGTGAACAAGCTGTGGACTA 1636
Qy 527 rProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIleTh 547
Db 1637 CCCTCCCGTGGCCCAATGACGAGCCCAACCAAGTGTATCACCCTGCTCAGAACTCCATCAC 1696
Qy 547 rLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleValLeuTyTrpSerLeuG1 567
Db 1697 CCTCTTTGGAAACACGAGCAGCGATGATGACCGGATACACAGCTATGAGTGTGCTCAG 1756
Qy 567 yProGlySerGluGlyLysHisValValMetGlnGlyValGlnThrProTyTrpLeuHisLe 587
Db 1757 CCCGAGCAGCAAGGGAAGTGGTGGAGATGACGGAGTTAGAACGCCAGCCCTGCAGCT 1816
Qy 587 uSerAlaMetGlnGluGlyAspTyTrpPheGlnLeuLysValThrAspSerSerArgL 607
Db 1817 GTCCGCAATGCAAGAGGAGACTATACCTACAGCTCAGCTGACTGACACCGCAGGACA 1876
Qy 607 nGlnSerThrAla**ValThrValIleValGlnProGluAsnAsnArgProProValAl 627
Db 1877 ACAGGCCACCCCAAGTACTGTGATGTGACGCTGAGAACCAACAGCCCTCTCAGGC 1936
Qy 627 aValAlaGlyProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGlySe 647
Db 1937 AGATGACGGCCACACAAAGAGCTGACCTGCCGTGGACAGCAACCCCTGGACGGCAG 1996
Qy 647 rSerSerSerAspAspHisGlyIleValPheTyTrpHisThrGluHisValArgGlyProSe 667
Db 1997 CAAGAGCAGACATGACACAGAGAGCTGCTCTTACCTTTGGAGCAGAGTCCGGGACCTGA 2056
Qy 667 rAlaValGluMetGluAsnIleAspLysAlaIleAlaThrValThrGlyLeuGlnValG1 687
Db 2057 CGGGGTGAGCTGGAGAAATGCCAAGCAGCTGTCGCCACTGTGACTGTGCTGCAAGTCGG 2116
Qy 687 yThrTyHisPheArgLeuThrValLysAspGlnGlnGlyLeuSerSerThrThrLe 707
Db 2117 GACTTATGATATCCCTTGACTGTCTCAAGATGAGAGGAACTACAGAGCCAGAGCTCCGT 2176
Qy 707 uThrVal 709
Db 2177 TAATGTC 2183

RESULT 11
AUI30115
LOCUS AUI30115 NT2RP3 Homo sapiens cDNA clone NT2RP3000279 5', mRNA
DEFINITION AUI30115 842 bp mRNA linear EST 01-AUG-2002
sequence.
ACCESSION AUI30115
VERSION AUI30115.1 GI:10990469
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 842)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
```

```
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1. 842
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3000279"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RP3"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

FEATURES
source
1. 842

ORIGIN
Alignment Scores:
Pred. No.: 1,01e-107 Length: 842
Score: 1312.00 Matches: 267
Percent Similarity: 96.43% Conservative: 3
Best Local Similarity: 95.36% Mismatches: 6
Query Match: 23.51% Indels: 5
DB: 1 Gaps: 0

US-10-764-390-3 (1-1072) x AUI30115 (1-842)
Qy 347 LeuProAspAsnGluValGluLeuLysAlaPheValAlaProAlaProValGluThr 366
Db 3 TTATCCCGACAAATGAAGTTGAACCTGAAGCCCTTTGCGCCAGCCACCTGTAGAACA 62
Qy 367 ThrTyHisThrGluTrpAsnLeuIleSerHisProThrAspTyTrpGlnGlyLys 386
Db 63 ACCTACAACTATGAAATGGAATTTAATAGCCACCCACAGACTACCAAGGTGAATAA 122
Qy 387 GlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyValPheLys 406
Db 123 CAAGGACACAAAGCAACTCTTAACCTCTCAATGTCGGTCCGACCTTATGTCTTCAA 182
Qy 407 ValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValLysPro 426
Db 183 GTCACTGTTCTAGTGAACCGCTTTGGAGAAGGATTTGTCAATGTCTCACTGTTAAGCCT 242
Qy 427 AlaArgValAsnLeuProProValAlaValValSerProGlnLeuGluLeuThr 446
Db 243 GCCAAGAGTCAACCTGCCACCTGTAGCAGTTGTTCTCCCACTGCAAGAGCTCACT 302
Qy 447 LeuProLeuThrSerAlaLeuLeuAspGlySerGlnSerThrAspAspThrGluLeuVal 466
Db 303 TTGCTTTGAGCTGAGCCCTCATTTGATGGCAGCCCAAGTACAGATGATGAAATAGTG 362
Qy 467 SerTyHisTrpGluGluIleAsnGlyProPheIleGluGlySerValAspSer 486
Db 363 AGTTATCATTTGGGAAGAAATAAACGGGCCCTTCATAGAAGAGAAGACTTCAGTTGACTCT 422
Qy 487 ProValLeuArgLeuSerAsnLeuAspProGlyAsnTyTrpSerPheArgLeuThrValThr 506
Db 423 CCCGCTCTTACGCTTGTCTAACCTTGTGCTGTAACATATAGTTTCAGGTTGACTGTACA 482
Qy 507 AspSerAspGlyAlaThrAsnSerThrAlaAlaLeuLeuValAsnAsnAlaValAsp 526
Db 483 GACTCGGACGGAGCACTAATCTCAACTGACGCCCTTAATAGTGACAACTGCTGTGAC 542
Qy 527 TyrProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIle 546
Db 543 TACCCACAGTTGCTTAATGCGAGGACCAAAATCACACCATAACTTTGCCCCCAAACTCCATC 602
Qy 547 ThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleValLeuTyTrpSerLeu 566
Db 603 ACTTTGAATGAAACCAAGCAGCAGTACCAAGATTCACAGATTTGCTCTATAGAGTGGTCCCTG 662
Qy 567 GlyProGlySerGluGlyLysHisValValMetGlnGlyValGlnThrProTyTrpLeuHis 586
Db 663 GGTCTGGAGTGANGGCAACATGTGTGTCATGAGGGAGT-CAGACGCCATACCTTCAT 721
```


QY 587 LeuSerAlaMetGlnGluGlyAspTyrThrPheGln-LeuLysValThrAsp-SerSera 606
 |||||
 Db 722 TTATCTGCAATCAGGAGGAGATATCATCTTACCTTGTGAGGTGACAGAACTCTTAA 781
 |||||
 QY 606 rgGlnGlnSerThrAla-***ValThr-ValLleValGlnProGluuAsnAsn 622
 |||||
 Db 782 GGCACACAGTCTACTGCTGGGGGTGACTGNGAATGGTCCAGGCTGAAACCAT 833
 |||||

RESULT 12
 CB722905 710 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-GH0-cef-1-19-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
 DEFINITION IMAGE:6838844 5', mRNA sequence.

ACCESSION CB722905
 VERSION CB722905.1 GI:29780047

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 710)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

<http://genome.uiowa.edu/distribution/mousefl.html>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .710

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6838844"

/tissue_type="Whole brain"

/dev_stage="1, 5, and 15 days newborn"

/lab_host="NIH BMAP GH0"

/clone_lib="NIH BMAP GH0"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bernaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CGAATGAT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 3.07e-84 Length: 710
 Score: 1051.00 Matches: 203
 Percent Similarity: 92.31% Conservative: 13
 Best Local Similarity: 86.75% Mismatches: 18
 Query Match: 18.84% Indels: 0
 DB: 6 Gaps: 0

US-10-764-390-3 (1-1072) x CB722905 (1-710)

QY 575 ValValMetGlnGlyValGlnThrProTyrThrLeuHisLeuSerAlaMetGlnGluGlyAsp 594
 |||||
 Db 4 ATGCTCATGAGGAGGCGCAGACCCCATACCTTCTACTTGTCTGAGCTCGAGGAGGAGAG 63
 |||||
 QY 595 TyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla-***ValThr 614
 |||||
 Db 64 TACACATTTACGCTGATGGTGCAGCATTCCTCGGACAGCAATCCACCGCTTTGGTCGCT 123
 |||||
 QY 615 ValLleValGlnProGluuAsnAsnAsnProProValAlaValAlaValAlaGlyProAspLysGlu 634
 |||||
 Db 124 GTGACTGTCCAGGCTGAGAACACAGCGCTCCGTGGCCGTGGCAGGCCCCGACAAAGGAG 183
 |||||
 QY 635 LeuLlePheProValGluSerAlaThrLeuAspGlySerSerSerSerSerSerSerSerSer 654
 |||||
 Db 184 CTGGTTTCCCGGTGCAGAGTGCTACACTGATGGGACCGGAGCAGTGACGATCACGGC 243
 |||||
 QY 655 IleValPheTyrHisTrrpGluHisValArgGlyProSerAlaValGluMetGluAsnLe 674
 |||||
 Db 244 ATTGTCTGTCTATCACTGGGAGCACATCAGAGGCCCCAGTCGAGTCGAGATGGAGAATGTG 303
 |||||
 QY 675 AspLysAlaLleAlaThrValThrGlyLeuGlnValGlyThrTyrThrHisPheArgLeuThr 694
 |||||
 Db 304 GACAAAGCCATAGCCACTGTCCCGGCCCTTCAGGTGGGTATCTACCACTTCGGGTAAACC 363
 |||||
 QY 695 VallysAspGlnGlnGlyLeuSerSerThrThrLeuThrValAlaValLysLysGlu 714
 |||||
 Db 364 GTGAGACACACAGAGGCTCTGAGCAGCAGCTCCACCCTCACTGTGGCTGTGAAGAGGAA 423
 |||||
 QY 715 AsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAsn 734
 |||||
 Db 424 AATAACAGCCCTCCAGAGCCAGCTGTGGTGGCAGACATGTTCTTATACTTCCCAATAAT 483
 |||||
 QY 735 SerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrLeuTrp 754
 |||||
 Db 484 TCCATTACTTTGGATGGTTCAGAGGTCTACTGATGACCGAGGAATGTGTCTCTATCTGG 543
 |||||
 QY 755 IleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerVal 774
 |||||
 Db 544 ATCCGGGATGCCAGAGTCCAGCTGCCGAGATGTCAATGGAGGCTCTGACCACAGGGCA 603
 |||||
 QY 775 AlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThrAsp 794
 |||||
 Db 604 GCTCTGCAGCTCAACAATCTGTTGGAGGCGTCTACACTTTTCACTTGTCTAGTCACTGAC 663
 |||||
 QY 795 SerGlnGlyAlaSerAspThrAspThrAlaThrValGluVal 808
 |||||
 Db 664 AGTCAGGGGGCTCGGACTCAGACGCGGCCACTGTGGAGGTG 705
 |||||

RESULT 13

LOCUS AL044313

DEFINITION DKFZp3400729_r1 434 (synonym: htes3) Homo sapiens cDNA clone
 DKFZp3400729 5', mRNA sequence.

ACCESSION AL044313

VERSION AL044313.1 GI:5432536

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 597)

AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Koehrer, et al.)

JOURNAL Unpublished (1999)

COMMENT Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by BMFZ (Biomedical Research Center at the Charite,

Berlin/Germany) within the cDNA sequencing consortium of the German

Genome Project.
No. 81 sequence available.
This clone (DKFp43400729) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
Location/Qualifiers
1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFp43400729"
/issue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Alignment Scores:
Pred. No.: 5,18e-82 Length: 597
Score: 1025.00 Matches: 197
Percent Similarity: 98.99% Conservative: 0
Best Local Similarity: 98.99% Mismatches: 1
Query Match: 18.37% Indels: 1
DB: 1 Gaps: 0

US-10-764-390-3 (1-1072) x AL044313 (1-597)

Qy 467 SerTyrHisTrpGluGluLeuAsnGlyProPheLeuGluLysThrSerValAspSer 486
Db 3 AGTTATCATTCGGGAAGAAATAAACGGCCCTTCATAGAAGAGAGACTTCAGTTGACTCT 62

Qy 487 ProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSerPheArgLeuThrValThr 506
Db 63 CCCGCTTACGCTTGCTCTAACCTTCATCCTGGTAACATAGTTTCAGGTTGACTGTACA 122

Qy 507 AspSerPheGlyAlaThrAsnSerThrThrAlaAlaLeuLeuValAsnAsnAlaValAsp 526
Db 123 GACTCGGAGGAGCCCACTAATCTACAACTGAGCCCTTAATAGTGAACAATGCTGTGGAC 182

Qy 527 TyrProProValAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIle 546
Db 183 TACCACCACTGTTGCTAATGAGGAGCAATACACCACTAATCTTGCCCAAACTCCATC 242

Qy 547 ThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleValLeuTyrGluTrpSerLeu 566
Db 243 ACTTTGAATGAAACACGAGCAGTGACATCACCAGATTGCTCTATAGTGTGCTCCTG 302

Qy 567 GlyProGlySerGluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeuHis 586
Db 303 GGTCTGGAGTGAGGGCAACATGTGCTCATGACGAGGAGTACAGACCCATACCTTCAT 362

Qy 587 LeuSerAlaMetGlnGluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArg 606
Db 363 TTAICTGCAATGACGAGGAGATTAATACATTTGAGCTGAGGAGTACAGATTCCTCAAGG 422

Qy 607 GlnGlnSerThrAla**ValThrValIleValGlnProGluAsnAsnArgProProVal 626
Db 423 CAACAGTCTACTGCTGTGCTGACTGTGATTGTCCAGCCTGAAACAAATAGACCTCCAGTG 482

Qy 627 AlaValAlaGlyProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGly 646
Db 483 GCTGTGGCGGCGCTGATAAAGAGCTGATCTTCCAGTGGAAAGTGTCTACCTGGATGGG 542

Qy 647 SerSerSerAspAspHisGlyIleValPheTyrHisTrp-GluHisValArg 664
Db 543 AGCAGCAGCGATGACACGGCAATTGTCTTACCACTGGGAGGACGTCAGA 597

RESULT 14
BP229849 598 bp mRNA linear EST 15-SEP-2004
LOCUS BP229849 Sugano cDNA library, fetal brain Homo sapiens cDNA clone
DEFINITION FBR04757, mRNA sequence.

ACCESSION BP229849
VERSION BP229849.1 GI:52102759
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Location/Qualifiers
1 (bases 1 to 598)
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FBR04757"
/issue_type="brain"
/dev_stage="fetal"
/clone_lib="Sugano cDNA library, fetal brain"

ORIGIN
Alignment Scores:
Pred. No.: 1.17e-79 Length: 598
Score: 999.00 Matches: 183
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 98.92% Mismatches: 0
Query Match: 17.90% Indels: 0
DB: 5 Gaps: 0

US-10-764-390-3 (1-1072) x BP229849 (1-598)

Qy 888 AsnLeuHisMetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuArg 907
Db 11 AATCTGCACATGCGCTCTCAAGAGAGAGCCGACTTCTTGTCTTTCAAGGCTCTGAGG 70

Qy 908 ValAspThrAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuThr 927
Db 71 GTTGATACAGCAGGTGCTCTTCTGAAGTGTCTGCGCATGTGTCACTGCGACCCCTCACA 130

Qy 928 LysArgCysIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAsp 947
Db 131 AAGCGTGCATTTGCTCTCACTTATGATGAGAGAACCTTATACGCGTTATATCTGGAT 190

Qy 948 GlyGluSerAsnCysGluTrpSerIlePheTyrValThrValLeuAlaPheThrLeuIle 967
Db 191 GGAGAGAGCAACTGTGAGTGGAGTATATCTATGTGACAGTGTGGCTTTTACTCTTATT 250

Qy 968 ValLeuThrGlyGlyPheThrTrpLeuCysLysCysCysLysArgGlnLysArgThr 987
Db 251 GTGCTAACAGAGGTTTCACTTGGCTTTGCACTCTGCTGCTCAAAAGACAAAAGGACT 310

Qy 988 LysIleArgLysLysThrLysTyrThrIleLeuAspAsnMetAspGluGlnGluArgMet 1007
Db 311 AAAATCAGGAAAAAACAAGTACCACTCTGGATACATGGATGAACAGGAAGAATG 370

Qy 1008 GluLeuArgProLysTyrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMet 1027
Db 371 GAATCGAGCCCAATATGTTATCAAGCACCGAGCAGACACAACTCCAGCTGTATG 430

Qy 1028 ValSerGluSerGluPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGlu 1047
Db 431 GTATCCGAGTCTGAGTTTGACAGTGACCCAGGACACAATCTTCAGCCGAGAGAGAGAG 490

Qy 1048 ArgGlyAsnProLysValSerMetAsnGlySerIleArgAsnGlyAlaSerPheSerTyr 1067

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 12, 2005, 10:29:21 ; Search time 1169 Seconds

(without alignments)

5428.540 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 5580

Sequence: 1 MAPPTGVLSLLLVITAGC.....VSMNGSRNGASFSYCSKDR 1072

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US10764390/runat_12102005_110203_22319/app_query.fasta_1.1223
-DB=N_Geneseq_16Dec04 -QWMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10764390@cgn 1 1 708 @runat_12102005_110203_22319 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*

1:	geneseqn1980s:*
2:	geneseqn1900s:*
3:	geneseqn2000s:*
4:	geneseqn2001as:*
5:	geneseqn2001bs:*
6:	geneseqn2002as:*
7:	geneseqn2002bs:*
8:	geneseqn2003as:*
9:	geneseqn2003bs:*
10:	geneseqn2003cs:*
11:	geneseqn2003ds:*
12:	geneseqn2004as:*
13:	geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5574	99.9	6791	13 ADR00593	Adr00593 Human 254
2	5574	99.9	6791	13 ADR00591	Adr00591 Human 254
3	5546.5	99.4	6797	13 ADR00860	Adr00860 Human 254
4	5546.5	99.4	6797	13 ADR00859	Adr00859 Human 254
5	5527	99.1	6991	13 ADR00595	Adr00595 Human 254

6	5527	99.1	6991	13	ADR000858	Adr00858 Human 254	
7	5358	96.0	3808	10	ADG15055	Adg15055 Human SEC	
8	2435.5	43.6	3147	12	ADM36226	Adm36226 Murine SS	
9	2404	43.1	3846	10	ADI21796	Adi21796 Novel hum	
10	2404	43.1	3908	4	AAD11067	Aad11067 Human sec	
11	2385	42.7	4001	10	ADF42453	Adf42453 Human PP7	
12	2293.5	41.1	3918	5	AAH89908	Aah89908 Human bon	
13	2278	40.8	3628	10	ADG15051	Adg15051 Human SEC	
14	2184.5	39.1	2796	4	AAH18563	Aah18563 Human cDN	
15	1817.5	32.6	2501	2	AAV58761	Aav58761 Human sec	
16	1451.5	26.0	3345	4	ABL13095	Ab113095 Drosophil	
17	1443.5	25.9	2223	4	AAH15983	Aah15983 Human cDN	
18	1314	23.5	2423	5	AAH73181	Aah73181 DNA encod	
19	1280.5	22.9	5775	4	ABL13094	Ab113094 Drosophil	
20	845	15.1	772	4	AAH06733	Aah06733 Human cDN	
21	823	14.7	818	8	ACA56915	Aca56915 Human adi	
22	812.5	14.6	1733	5	AAH90021	Aah90021 Human bon	
23	681	12.2	580	2	AAV89845	Aav89845 EST clone	
24	640	11.5	578	2	AAV89613	Aav89613 EST clone	
C	25	477	8.5	284	13	ADR00590	Adr00590 Human 254
C	26	437	7.8	982	4	AAI89864	Aai89864 Human pol
C	27	403	7.2	126990	12	ADP13332	Adp13332 Renal cel
C	28	397	7.1	1420	3	AAC69580	Aac69580 Human sec
C	29	332	5.9	313	6	ABN22934	Abn22934 Human ORF
C	30	241.5	4.3	3612	5	ACA25573	Aca25573 Prokaryot
C	31	235.5	4.2	20320	5	ABA21464	Aba21464 Human ner
C	32	225	4.0	103052	13	ADQ89963	Adq89963 Antagonis
C	33	224	4.0	81940	4	AAK05390	Aas05390 Human tit
C	34	224	4.0	81940	6	ABK64829	Abk64829 Human ben
C	35	224	4.0	81940	12	ADQ17315	Adq17315 Human sof
C	36	222.5	4.0	6614	10	ADC30653	Adc30653 Human nov
C	37	216.5	3.9	110000	12	ADN46845	Adn46845 Human nov
C	38	216.5	3.9	110000	12	ADN47591	Adn47591 Human nov
C	39	216.5	3.9	110000	12	ADN46123	Adn46123 Human nov
C	40	216.5	3.9	110000	12	ADN47209	Adn47209 Human nov
C	41	216.5	3.9	110000	12	ADN46464	Adn46464 Human nov
C	42	216.5	3.9	110000	12	ADN47960	Adn47960 Human nov
C	43	214.5	3.8	3687	8	ABZ77634	Abz77634 Nucleotid
C	44	214.5	3.8	3687	12	ADH36637	Adh36637 DNA encod
C	45	214	3.8	23545	13	ADQ89759	Adq89759 Antagonis

ALIGNMENTS

RESULT 1

ADRO0593

ID ADRO0593 standard; cDNA; 6791 BP.

XX AC

XX ADRO0593;

XX

XX

XX 04-NOV-2004 (first entry)

XX

XX Human 254P1D6B v.2 encoding cDNA SEQ ID NO:4.

XX

XX 254P1D6B; small interfering RNA; siRNA; immune response;

XX KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;

XX KW 254P1D6B v.2; gene; ss; chromosome 6.

XX

XX Homo sapiens.

XX

XX

XX Key Location/Qualifiers

XX CDS 512..3730

XX

XX FT /*tag= a

XX

XX FT /product= "254P1D6B v.2"

XX

XX WO2004067716-A2.

XX

XX

XX 12-AUG-2004.

XX

XX 23-JAN-2004; 2004WO-US001965.

XX

XX 24-JAN-2003; 2003US-0442526P.

XX

XX

ALIGNMENTS

RESULT 1	ADR00593	standard; cDNA; 6791 BP.
ID	ADR00593	(first entry)
XX	ADR00593	
AC	ADR00593	
XX	04-NOV-2004	
DT	Human 254P1D6B v.2	encoding cDNA SEQ ID NO:4.
DE	254P1D6B	small interfering RNA; siRNA; immune response;
XX	254P1D6B	related protein; cytostatic; gene therapy; cancer; human;
KW	254P1D6B v.2	gene; ss; chromosome 6.
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FH	CDS	512..3730
FT		/*tag= a
FT		/product= "254P1D6B v.2"
XX	WO2004067716-A2.	
PN	12-AUG-2004.	
XX	23-JAN-2004;	2004WO-US001965.
XX	24-JAN-2003;	2003US-0442526P.
XX		

PA	(AGEN-) AGENSYS INC.	QY	81	SerCysProHisLysGluAsnGluProLysLysMetGlyProIleArgSerTyrLeu	100
XX					
PI	Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;	Db	752	AGCTGCCCCACAAAGAGAACTGTGAGCCCAAGAGATGGGCCCATCAGGTCTTATCTC	811
PI	Perez-Villar JJ, Paris M;				
XX		QY	101	ThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMet	120
DR	WPI: 2004-580991/56.	Db	812	ACTTTTGTGCTCCGGCCCTGTTTCAGAGGCTGCACAGCTGTGGACTATGGGACATGATG	871
DR	P-P8DB; ADR00594.				
XX		QY	121	LeuAsnArgGlySerProSerGlyLleTTPGlyAspSerProGluAspIleArgLysAsp	140
PT	New 254P1D6B siRNA composition comprising a double stranded siRNA that				
PT	corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B				
PT	protein or corresponds to a subsequence of the ORF, useful for detecting	Db	872	CTGAACAGGGGCTCCCCCTCGGGATCTCGGGGAGCTCACCTGAGGATATCAGAAAGGAC	931
PT	and treating cancer.				
XX		QY	141	Leu***PheLeuGlyLysAspTTPGlyLeuGluMetSerGluTyr***AspAspTyr	160
PS	Claim 10; SEQ ID NO 4; 345pp; English.	Db	932	TTGCCCTTTCTAGGCAAGATTGGGGCCCTAGAGGAGATGTCTGAGTACGCAGATGACTAC	991
XX		QY	161	ArgGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAla	180
CC	The present invention describes a 254P1D6B small interfering RNA (siRNA)				
CC	composition that comprises a double stranded siRNA that corresponds to				
CC	the nucleic acid open reading frame (ORF) sequence which encodes the				
CC	254P1D6B protein, or corresponds to a subsequence of the ORF, where the	Db	992	CGGGAGCTGGAGAGGACCTCTTGCAACCCAGTGGCAAGCAGAGGCCAGAGGGAGTGCC	1051
CC	double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous				
CC	nucleotides in length. Also described: (1) a composition that comprises,	QY	181	GluTyrThrAspTTPGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGly	200
CC	consists essentially of, or consists of a peptide of eight, nine, ten, or				
CC	eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID	Db	1052	GAGTACAGGACTGGGGCTTACTGCCGGGCGAGCGGGGGCTTCAACTCTCTCTGTGGA	1111
CC	NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in				
CC	any of the 42 lists of peptides, given in the specification, or a protein	QY	201	AspSerProAlaValProAlaGluThrGlnGlnAspProGluLeuLeuHisTyrLeuAsnGlu	220
CC	that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or				
CC	identical to an entire amino acid sequence of PI; (2) a polynucleotide	Db	1112	GACAGTCTCGGTGCCAGCGGAGACGACGACGAGGACCTCTGAGCTCCATTACTGTAATGAG	1171
CC	that encodes the protein; (3) a composition comprising a polynucleotide				
CC	that is a full complement of the polynucleotide described above; (4)	QY	221	SerAlaSerThrProAlaProLysLeuProGluArgSerValLeuLeuProLeuProThr	240
CC	generating a mammalian immune response directed to the protein of PI; (5)				
CC	detecting, in a sample, the presence of a 254P1D6B-related protein or a	Db	1172	TCGGCTTCAACCCCTCGCCCAAACTCCCTGAGAGAGTGTGTGCTTCCCTTGCCTGACT	1231
CC	254P1D6B-related polynucleotide; (6) a composition that modulates the				
CC	status of a cell that expresses a protein of PI; (7) a pharmaceutical	QY	241	ThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGlnGlnSer	260
CC	composition that comprises the composition described above in a human				
CC	unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)	Db	1232	ACTTCATCTTCAGAGAGGTGTGGAGAAAGAAAGGCTTCTCAGCTCCAGGNACAACTCC	1291
CC	a non-human transgenic animal that produces the antibody; (10) a	QY	261	SerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProProAlaSerLeu	280
CC	hybridoma that produces the antibody; (11) delivering a cytotoxic agent				
CC	or a diagnostic agent to a cell that expresses the protein of PI; and	Db	1292	AGCAACAGCTCTGGAAAGAGGTCTTAATGCTTCCCATAGTCTTCTCTCCGCAAGCCTG	1351
CC	(12) inhibiting growth, reproduction or survival of cancer cells that	QY	281	GluLeuSerSerValThrValGluLysSerProValLeuThrValThrProGlySerThr	300
CC	express the protein of PI. 254P1D6B has cytostatic activity, and can be				
CC	used in gene therapy. The compositions, molecules and methods are useful	Db	1352	GAGCTCAGCTCAGTCACCGTGAGAGAAAGCCAGTGTCTCACAGTCACCCCGGGAGTACA	1411
CC	for treating and detecting cancer. The present sequence encodes the human	QY	301	GluHisSerIleProThrProThrSerAlaAlaProSerGluSerThrProSerGlu	320
CC	254P1D6B v.2, which is used in the exemplification of the present				
CC	invention. The human 254P1D6B gene is located on chromosome 6p22.	Db	1412	GAGCAGCATCCCAACACCTCCCACTAGCGAGCCCTCTGAGTCCACCCCATCTGAG	1471
XX					
SQ	Sequence 6791 BP; 1812 A; 1542 C; 1684 G; 1753 T; 0 U; 0 Other;				
Alignment Scores:					
Pred. No.:	0	Length:	6791		
Score:	5574.00	Matches:	1069		
Percent Similarity:	99.72%	Conservative:	0		
Best Local Similarity:	99.72%	Mismatches:	3		
Query Match:	99.89%	Indels:	0		
DB:	13	Gaps:	0		
US-10-764-390-3 (1-1072) x ADR00593 (1-6791)					
QY	1	MetAlaProProThrGlyValLeuSerSerLeuLeuLeuValThrIleAlaGlyCys	20		
Db	512	ATGGCGCCCCACAGGTGTGCTCTTCAATTGCTGCTGTGTGAGCAATTCAGGTTGT	571		
QY	21	AlaArgLysGlnCysSerGluGlyArgThrTyrSerAsnAlaValIleSerProAsnLeu	40		
Db	572	GCCCGTAGCAGTGCAGCGGGGAGGACATATTCATGTCAGTCAATTCACCTAACTTG	631		
QY	41	GluThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAlaAla	60		
Db	632	GAACACCAACAGATCATGGGGTGTCTCACACCTTCCCTGTGTAGACTGCACGGCGCT	691		
QY	61	CysCysAspLeuSerSerCysAspLeuAlaTTPpPheGluGlyArgCysTyrLeuVal	80		
Db	692	TGCTGTGACCTGTCAGCTGTGACCTGGCCTGGTGGTTCGAGGGCGCTGCTACCTGGTG	751		

Db 1832 CAATGCAAGAGCTCATTGTCCTTTGACGTCAGCCCTCAFTGATGGGAGCCCAAGATACA 1891
Qy 461 AspAspThrGluIleValSerTyrHisTrpGluGluIleAenGlyProPheIleGluGlu 480
Db 1892 GATGATACTGAATAGTAGTATCAATGGGAAGAATAAAGCGGCCCTTCATAGAGAG 1951
Qy 481 LysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSer 500
Db 1952 AAGACTTCAGTTGACTCTCCGCTTATCGCTTGTCTAACCTTGATCTCGTGAATAGT 2011
Qy 501 PheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrAlaAlaLeuIle 520
Db 2012 TTCAGTTGACTGTTACAGACTCGGAGCGGCCCACTAACTCTCAACTGCGAGCCCTAATA 2071
Qy 521 ValAsnAsnAlaValAspTyrProProValAlaAsnAlaGlyProAsnHisThrIleThr 540
Db 2072 GTGAACAATGCTGTGGACTACCCACCACTTGTCTAATGCAGGACCAATCAACCATAACT 2131
Qy 541 LeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleVal 560
Db 2132 TTGCCCCAAAACCTCATCACTTTGAATGGAAACAGAGCAGTGACGATCAACGATTGTC 2191
Qy 561 LeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHisValMetGlnGlyVal 580
Db 2192 CTCTATGAGTGGTCCCTGGGTCCTGGGAGTGAGGCGCAACATGTGGTCATGCGGGAGTA 2251
Qy 581 GlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAspTyrThrPheGlnLeuLys 600
Db 2252 CAGAGCCCATACCTTCATTATCTGCAATGCAGGAAGGAGATTATACATTTACGTCGAG 2311
Qy 601 ValThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnProGlu 620
Db 2312 GTGACAGATTCTTCAAGGCAACAGCTACTGCTGTAGTGTACTGTGATTGTCCAGCCTGAA 2371
Qy 621 AsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProValGlu 640
Db 2372 AACAAATAGACTCCAGTGGCTGTGGCGGCCCTGATAAAGAGCTGATTTCCCGAGTGGAA 2431
Qy 641 SerAlaThrLeuAspGlySerSerSerSerAspAspHisGlyIleValPheTyrHisTrp 660
Db 2432 AGTCTACTCCCTGGATGGGAGCAGCAGCAGCAGTACCAACGCGATTGCTTACCACTGG 2491
Qy 661 GluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAlaThr 680
Db 2492 GAGCACGTCAGAGGCCCGAGTGCAGTGAGATGAAATAATTGACAAAGCAATAGCCACT 2551
Qy 681 ValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGly 700
Db 2552 GTGACTGGTCTCCAGGTGGGAGCCTACCACTTCCGTTTGACAGTGAAGACCAAGCGGGA 2611
Qy 701 LeuSerSerThrSerThrLeuThrValAlaValLysLysGluAsnAsnSerProProArg 720
Db 2612 CTGAGCAGCAGCTCCACCTCACTGTGGCTGTGAAGAGGAAATAATAGTCTCCAGAGA 2671
Qy 721 AlaArgAlaGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAspGly 740
Db 2672 GCCCGGGCTGTGGCAGACATGTTCTTGTGCTTCCCAATAATTCCATTACTTTGGATGGT 2731
Qy 741 SerArgSerThrAspAspGlnArgIleValSerTyrLeuTrpIleArgAspGlyGlnSer 760
Db 2732 TCAAGGTCTACTGATGACCAAGAATAATTGTGCTTCTGTGGATCCGGGATGGCCAGAGT 2791
Qy 761 ProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsn 780
Db 2792 CCAGCAGCTGGAGATGTATCATGGCTCTGACCACAGTGTGGCTCTGACCTTACGAAT 2851
Qy 781 LeuValGluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAsp 800
Db 2852 CTGGTGGAGGGGTGTACACTTTCCACTTTCAGAGTCCAGACAGTCAGGGGGCCTCGGAC 2911
Qy 801 ThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValGluLeu 820

Db 2912 ACAGACACTCCCACTGTGGAGTGCAGCCAGACCCCTAGGAAGAGTGGCTGGTGGAGCTG 2971
Qy 821 ThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuValArgGln 840
Db 2972 ACCCTGCAGGTGGTGTGGGAGCTGACAGAGCAGCGGAAGACACCCCTTGTGAGGCAG 3031
Qy 841 LeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLysIleAArgAlaHis 860
Db 3032 CTGGCTGTGTCTGTAACGTCGTGGACTCGGACATTAAAGGTCCAGAAAGATTCGGGCCAC 3091
Qy 861 SerAspLeuSerThrValIleValPheTyrValGlnSerArgProProPheLysValLeu 880
Db 3092 TCGGATCTCAGACCGGTGATGTGTTTATGTACAGAGCAGCGCCCTTTCAAGGTTCTC 3151
Qy 881 LysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAspPhe 900
Db 3152 AAAGCTGCTGAAGTGGCCGAAATCTGCACATGCGGCTCTCAAAGGAGAAAGCTGACTTC 3211
Qy 901 LeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGlyHis 920
Db 3212 TTGCTTTTCAAGGTCTTGGAGGTGTGATACAGCAGGTTGCCCTTCTGAAGTGTCTTGGCCAT 3271
Qy 921 GlyHisCysAspProLeuThrLysArgCysIleCysSerHisLeuTrpMetGluAsnLeu 940
Db 3272 GGTCACTGCCACCCCTCACAAAGCGCTGCATTTGCTCTCACTTATGATGGAGAACTT 3331
Qy 941 IleGlnArgTyrIleTrpAspGlyGluSerAsnCysGluTrpSerIlePheTyrValThr 960
Db 3332 ATACAGCGTTATATCTGGGATGGAGAGCACTGTGAGTGGAGTATATTCTATGTGACA 3391
Qy 961 ValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCysCys 980
Db 3392 GTGTGGCTTTTACTCTTATTGTGCTAACAGGAGGTTTCACTTGGCTTTGCATCTGCTGC 3451
Qy 981 CysLysArgGlnLysArgThrLysIleArgLysLysThrLysTrpThrIleLeuAspAsn 1000
Db 3452 TGCANAAGACAAAAGAGCTAAATCAGNAAAAAACAAGTACACCATCTCGATAAC 3511
Qy 1001 MetAspGlnGluArgMetGluLeuArgProLysTyrGlyIleLysHisArgSerThr 1020
Db 3512 ATGGATGAACAGGAAGAAATGGAACTGAGGCCCAATATGGTATCAAGCACCAGACACA 3571
Qy 1021 GluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIle 1040
Db 3572 GAGCACAACTCCAGCCTGATGGTATCCGAGTCTGAGTTTGCAGCTGACCGAGACACAATC 3631
Qy 1041 PheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerIleArg 1060
Db 3632 TTCAGCCGAGAAAAGATGGAGAGAGGGAATCCAAAGGTTTCCATGATGGTTCATCAGA 3691
Qy 1061 AsnGlyAlaSerPheSerTyrCysSerLysAspArg 1072
Db 3692 AATGGAGCTTCCCTCAGTTATTGCTCAAGGACAGA 3727

RESULT 2
ADRD00591
ID ADR00591 standard; cDNA; 6791 BP.
XX XX
AC ADR00591;
XX XX
DT 04-NOV-2004 (first entry)
XX XX
DE Human 254P1D6B v.1 clone LCP-3 encoding cDNA SEQ ID NO:2.
KW 254P1D6B; small interfering RNA; siRNA; immune response;
KW 254P1D6B-related protein; cytosolic; gene therapy; cancer; human;
KW 254P1D6B v.1 clone LCP-3; gene; ss; chromosome 6.
OS Homo sapiens.
XX XX
FH Key Location/Qualifiers
FT CDS 512..3730
FT /*tag= a

Qy 381 TyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeuAenLeuSerGlnLeuSerVal 400
Db 1652 TACCAAGGTGAATATAAACAAGGACACAAAGCAAACTCTTAACCTCTCTCAATTGTCGGTC 1711
Qy 401 GlyLeuTyrValPheLysValThrValSerSerGluAenAlaPheGlyGluGlyPheVal 420
Db 1712 GGACTTATGCTTCAAAAGTCACTGTTCTTAGTGAAACGGCTTTGGNAGAGGATTTGTC 1771
Qy 421 AsnValThrValLysProAlaArgValAenLeuProProValAlaValAenLeuSerPro 440
Db 1772 AATGTCACGTAAAGCCTGCCAGAAGAGTCAACCTGCCACCTGTAGCAGTGTGTTCTCC 1831
Qy 441 GlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThr 460
Db 1832 CAATGCAAGAGCTCACTTTCCTTGAAGTCAAGCTGAGTGTGAGTGGCAGCAAGTACA 1891
Qy 461 AspAspThrGluIleValSerTyrHisTrpGluGluLeuAenGlyProPheIleGluGlu 480
Db 1892 GATGATACTGAAATAGTGAGTTATCAATTTGGGAAGAAATAAAGCGGCCCTTCATAGAAAG 1951
Qy 481 LysThrSerValAspSerProValLeuArgLeuSerAenLeuAspProGlyAsnTyrSer 500
Db 1952 AAGACTTCAGTTGACTCTCCCGTCTTACCGTTGTCTAACCTTTGATCTGGTAACTATAGT 2011
Qy 501 PheArgLeuThrValThrAspSerAspGlyAlaThrAenSerThrThrAlaAlaLeuIle 520
Db 2012 TTCAAGTTGACTGTTTACAGACTCGGAGGAGCCCACTAACTCTACAACCTGCAGGCCCTAATA 2071
Qy 521 ValAenAenAlaValAspTyrProProValAlaAenAlaGlyProAenHisThrIleThr 540
Db 2072 GTGAACAACTGCTGGAGTACCCACCACTTCTAATGAGGAGCAAAATACACCACTAACT 2131
Qy 541 LeuProGlnAenSerIleThrLeuAenGlyAsnGlnSerSerAspAspHisGlnIleVal 560
Db 2132 TTGCCCCAAACTCCATCACTTTGAATGGAACACAGAGCAGTGACGATCACAGATTGTC 2191
Qy 561 LeuTyrGluTrpSerLeuGlyProGlySerGlnGlyLysHisValMetGlnGlyVal 580
Db 2192 CTCTATGAGTGGTCCCTGGGTCCTGGGAGTGAGGCAAAACATGTGTGTCATGAGGAGTA 2251
Qy 581 GlnThrProTyrLeuHisLeuSerAlaMetGlnGlnGlyAspTyrThrPheGlnLeuLys 600
Db 2252 CAGACGCCATACCTTCATTATCTGCAATGCAAGAGGAGATATATACATTTCACTGAAG 2311
Qy 601 ValThrAspSerSerArgGlnGlnSerThrAla**ValThrValIleValGlnProGlu 620
Db 2312 GTGACAGATTCTTCAAGGCAACAGTCTACTGCTGTGTGACTGTGTGTTGCCAGCTGAA 2371
Qy 621 AsnAenArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProValGlu 640
Db 2372 AACAAATAGACCTCCAGTGGCTGTGGCGGCCCTGATAAAGAGCTGATCTTCCAGTGGAA 2431
Qy 641 SerAlaThrLeuAspGlySerSerSerAspAspHisGlyIleValPheTyrHisTrp 660
Db 2432 AGTCTACCTCCGTGATGGAGCAGCAGCAGCATGACCAACGCGCATTTGCTTACCACTGG 2491
Qy 661 GluHisValArgGlyProSerAlaValGlnMetGluAenIleAspLysAlaIleAlaThr 680
Db 2492 GAGCAGCTCAGAGGCCCACTGCTGAGTGGAGATGAAATATATGACAAAGCAATAGCCACT 2551
Qy 681 ValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGly 700
Db 2552 GTGACTGGTCTCCAGGTGGGACCTACCCTTCCGTTTGTGACAGTGAAGACCAAGGGA 2611
Qy 701 LeuSerSerThrSerThrLeuThrValAlaValLysLysGluAenAenSerProProArg 720
Db 2612 CTGAGCAGCAGCTCCACCTCACTGTGGCTGTGAAGAGGAAATAATAGTCTCTCCAGA 2671
Qy 721 AlaArgAlaGlyGlyArgHisValLeuValLeuProAenAenSerIleThrLeuAspGly 740
Db 2672 GCCCGGGCTGGTGACAGCATGTTCTGTGCTTCCCAATAATTCATTTACTTTGGATGGT 2731

Qy 741 SerArgSerThrAspAspGlnArgIleValSerTyrLeuTrpIleArgAspGlyGlnSer 760
Db 2732 TCAAGGCTTACTGATGACCAAGAAATTTGTCTTATCTGTGATCCGGGATGGCCAGAGT 2791
Qy 761 ProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsn 780
Db 2792 CCAGCAGCTGGAGATGTCATCGATGGCTCTGACCACACAGTGTGGCTCTGCAGCTTACGAAT 2851
Qy 781 LeuValGluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAsp 800
Db 2852 CTGTGGAGGGGTGTACACTTTCCTTCCAGTTCACCGACAGTCAAGGGGGCTTCGAC 2911
Qy 801 ThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValGluLeu 820
Db 2912 ACAGACACTGCTGAGTGGGAGTGCAGCCAGCTTCCAGGAGTGGGCTCTGGTGAGCTG 2971
Qy 821 ThrLeuGlnValGlyValGlnLeuThrGluGlnArgLysAspThrLeuValArgGln 840
Db 2972 ACCCTGCAGGTTGGTGTGGGAGCTGCAGAGCAGCGGAGGACACCTTGTGAGGAG 3031
Qy 841 LeuAlaValLeuLeuAenValLeuAspSerAspIleLysValGlnLysIleArgAlaHis 860
Db 3032 CTGGCTGTGCTGTAACGCTGCTGACACTCGGACATTAAGGTCCAGAGATTCGGGCCAC 3091
Qy 861 SerAspLeuSerThrValIleValPheTyrValGlnSerArgProProPheLysValLeu 880
Db 3092 TCGGATCTCAGCACCGGTGATGTTTATGTACAGAGCAGCGGCCCTTTCAAGGTTCTC 3151
Qy 881 LysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAspPhe 900
Db 3152 AAAGCTGCTGAGTGGGCCCGAAATCTGCACATGCGGCTCTCAAGGAGAAAGCTGACTTC 3211
Qy 901 LeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGlyHis 920
Db 3212 TTGCTTTTCAAGGCTTTGAGGGTTGATACAGAGGTTGCCCTTCTGAAGTGTCTCGCCAT 3271
Qy 921 GlyHisCysAspProLeuThrLysArgCysIleCysSerHisLeuTrpMetGluAenLeu 940
Db 3272 GGTCACTCGCAACCCCTCACAAAGCGCTGCATTTGCTCTCACTTATGGATGGAGAACCTT 3331
Qy 941 IleGlnArgTyrIleTrpAspGlyGluSerAsnCysGluTrpSerIlePheTyrValThr 960
Db 3332 ATACAGGTTATATCTGGGATGGAGAGCACTGTGAGTGGAGTATATTTCTATGTGACA 3391
Qy 961 ValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCysCys 980
Db 3392 GTGTTGGCTTTACTTCTTATTGCTAACAGAGGTTTCACTTGGCTTTGCATCTGCTGC 3451
Qy 981 CysLysArgGlnLysArgThrLysIleArgLysLysThrLysTyrThrIleLeuAspAsn 1000
Db 3452 TGCAAAAGACAAAAGAGACTAAATCAGGAAAAAACAAGTAGTACCATCTCTGGATAAC 3511
Qy 1001 MetAspGluGlnGluArgMetGluLeuArgProLysTyrGlyIleLysHisArgSerThr 1020
Db 3512 ATGATGAACAGGAAAGATGGAATGAGGCCCAATATGTTATCAAGCACCGGAGCACA 3571
Qy 1021 GluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIle 1040
Db 3572 GAGCACAACCTCCAGCTGATGGTATCCGAGTCTGAGTTTGACAGTGCAGCAGCAGCAATC 3631
Qy 1041 PheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerIleArg 1060
Db 3632 TTCAGCCGAGAAAAGATGGAGAGGGAATCCAAAGGTTTCCATGAATGGTTCCATCAGA 3691
Qy 1061 AsnGlyAlaSerPheSerTyrCysSerLysAspArg 1072
Db 3692 AATGAGCTTCTCTTCAAGTATTTGCTCAAGGACAGA 3727
RESULT 3
ID ADR00860
XX ADR00860 standard; DNA; 6797 BP.
AC ADR00860;

XX	04-NOV-2004	(first entry)	US-10-764-390-3 (1-1072) x ADR00860 (1-6797)
DT	Human 254P1D6B v.3	nucleotide sequence SEQ ID NO:271.	
DE	254P1D6B; small interfering RNA; siRNA; immune response;		
XX	254P1D6B-related protein; cytostatic; gene therapy; cancer; human;		
KW	254P1D6B v.3; chromosome 6; gene; ds.		
KW	Homo sapiens.		
XX	WO2004067716-A2.		
XX	12-AUG-2004.		
XX	23-JAN-2004; 2004WO-US001965.		
XX	24-JAN-2003; 2003US-0442526P.		
XX	(AGEN-) AGENSYS INC.		
XX	Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;		
PI	Perez-Villar JJ, Faris M;		
XX	WPI; 2004-580991/56.		
XX	New 254P1D6B siRNA composition comprising a double stranded siRNA that		
PT	corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B		
PT	protein or corresponds to a subsequence of the ORF, useful for detecting		
PT	and treating cancer.		
XX	Example 5; SEQ ID NO 271; 345pp; English.		
XX	The present invention describes a 254P1D6B small interfering RNA (siRNA)		
CC	composition that comprises a double stranded siRNA that corresponds to		
CC	the nucleic acid open reading frame (ORF) sequence which encodes the		
CC	254P1D6B protein, or corresponds to a subsequence of the ORF, where the		
CC	double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous		
CC	nucleotides in length. Also described: (1) a composition that comprises,		
CC	consists essentially of, or consists of a peptide of eight, nine, ten, or		
CC	eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID		
CC	NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in		
CC	any of the 42 lists of peptides, given in the specification, or a protein		
CC	that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or		
CC	identical to an entire amino acid sequence of PI; (2) a polynucleotide		
CC	that encodes the protein; (3) a composition comprising a polynucleotide		
CC	that is a full complement of the polynucleotide described above; (4)		
CC	generating a mammalian immune response directed to the protein of PI; (5)		
CC	detecting, in a sample, the presence of a 254P1D6B-related protein or a		
CC	254P1D6B-related polynucleotide; (6) a composition that modulates the		
CC	status of a cell that expresses a protein of PI; (7) a pharmaceutical		
CC	composition that comprises the composition described above in a human		
CC	unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)		
CC	a non-human transgenic animal that produces the antibody; (10) a		
CC	hybridoma that produces the antibody; (11) delivering a cytotoxic agent		
CC	or a diagnostic agent to a cell that expresses the protein of PI; and		
CC	(12) inhibiting growth, reproduction or survival of cancer cells that		
CC	express the protein of PI. 254P1D6B has cytostatic activity, and can be		
CC	used in gene therapy. The compositions, molecules and methods are useful		
CC	for treating and detecting cancer. The present sequence represents the		
CC	human 254P1D6B v.3 nucleotide sequence, which is used in the		
CC	exemplification of the present invention. The human 254P1D6B gene is		
CC	located on chromosome 6p22.		
XX	Sequence 6797 BP; 1812 A; 1542 C; 1686 G; 1757 T; 0 U; 0 Other;		
SQ	Alignment Scores:		
	Pred. No.:	0	Length: 6797
	Score:	5546.50	Matches: 1069
	Percent Similarity:	99.53%	Conservative: 0
	Best Local Similarity:	99.53%	Mismatches: 3
	Query Match:	99.40%	Indels: 2
	DB:	13	Gaps: 1

QY	1	MetAlaProThrGlyValLeuSerSerLeuLeuValThrIleAla----	G1 19
Db	514	ATGGCGCCCCCAGAGGTGCTCTCTTCATTGTCTGGTGGGACAAATTCAGTTGG	573
QY	19	YCysAlaArgLysGlnCysSerGluGlyArgThrTy-SerAsnAlaValIleSerProAs	39
Db	574	TTGTGCCGTAAAGCAGTCAGCGAGGGGAGACATATTCCAATGCAGTCATTTCACCTAA	633
QY	39	nLeuGluThrThrArgIleMetArgValSerHisThrPheProValValAlaSerCysThrAl	59
Db	634	CTTGGAAACACACAGAAATCATGCGGGTCTCTCACACCTTCCTGCTGTAGACTGCACGCG	693
QY	59	aAlaCysCysAspLeuSerSerCysAspLeuAlaTTrpPheGluGlyArgCysThrIle	79
Db	694	CGCTTGTGTGACCTGTCCAGCTGTGACCTGGCTGGTGGTTCGAGGGCCGCTGTACTCT	753
QY	79	uValSerCysProHisLysGluAsnGluProLysLysMetGlyProIleArgSerTy	99
Db	754	GGTGAGCTGCCCCCAGAGAACTGTGAGCCCCCAAGAGATGGGCCCATCAGTCTTA	813
QY	99	rLeuThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMe	119
Db	814	TCTCACTTTTGTGCTCGGCTGTTCAGAGGCTGTGCACAGCTGTGTGGACTATGGGACAT	873
QY	119	tMetLeuAsnArgLysSerProSerGlyIleTTrpGlyAspSerProGluAspIleArgLys	139
Db	874	GATCTGAACAGGGGCTCCCTCGGGGATCTGGGGGACTCACCTGAGGATATCAGAAA	933
QY	139	sAspLeu***PheLeuGlyLysAspTrpGlyLeuGluMetSerGluTy***AspAs	159
Db	934	GGACTTGCCCTTTCTAGCCAAAGATTGGGGCTTAGAGGAGATGCTGTAGTACTCAGATGA	993
QY	159	pTyArgGluLeuGlyLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySe	179
Db	994	CTACCGGAGCTGGAGAGGACCTCTTGCAACCCAGTGGCAAGCAGGAGGCCAGAGGGAG	1053
QY	179	rAlaGluTyThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerVa	199
Db	1054	TGCCAGTACTCTGCGGTGCCAGCGGAGACGACGAGGACCCCTGAGCTCCATTACTGAA	1173
QY	199	lGlyAspSerProAlaValProAlaGluThrGlnAspProGluLeuHisTyLeuAs	219
Db	1114	TGGAGACAGTCTCTGCGGTGCCAGCGGAGACGACGAGGACCCCTGAGCTCCATTACTGAA	1173
QY	219	nGluSerAlaSerThrProAlaProLysLeuProGluArgSerValLeuLeuProLeuPr	239
Db	1174	TGAGTGGGCTTCAACCCCTGCCCCCAAACTCCCTGAGAGAAAGTGTGTGTCTTCCCTTGC	1233
QY	239	oThrThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGlnGlu	259
Db	1234	GACTACTCCATCTTCAGGAGAGGTGTTGGAGAAAGAAAGGCTTCTCAGCTCCAGAAACA	1293
QY	259	nSerSerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProAlaSe	279
Db	1294	ATCCAGCAACAGCTCTGGAAAGAGGTTCATATGCTTCCCATAGTCTTCCCTCCGCAAG	1353
QY	279	rLeuGluLeuSerSerValThrValGluLysSerProValLeuThrValThrProGlySe	299
Db	1354	CTTGAGCTCAGCTCAGTCAGTCACGTCGAGAAAGCCAGTGTCTCAGACTCACCCTGGGAG	1413
QY	299	rThrGluHisSerIleProThrProThrSerAlaAlaProSerGluSerThrProSe	319
Db	1414	TACAGGACAGCATCCCAACACCTCCCATAGCCAGCCCTCTGAGTCCACCCCATC	1473
QY	319	rGluLeuProIleSerProThrThrAlaProArgThrValLysGluLeuThrValSerAl	339
Db	1474	TGAGCTACCCATATCTCTACACCTGCTCCAGGACAGTGAAGAACTTACGGTATCGGC	1533
QY	339	aGlyAspAsnLeuIleThrLeuProAspAsnGluValGluLeuLysAlaPheValAl	359

Db 1534 TGGAGATTAACCTAATATTAACCTTACCAGCAATGAAGTTGAACCTGAAGGCTTTGTGTC 1593
Qy 359 aProAlaProValGluThrThrTyAsnTyxGluTrpAsnLeuLeuSerHisProTh 379
Db 1594 GCCAGCGCCACTGTAGAAAACAACCTAACAATATGAATGAATTAATAAGCCACCCAC 1653
Qy 379 rAspTyxGlnGlyGluIleValSerGlnGlyHisValSerGlnLeuLeuSerGlnLeuSe 399
Db 1654 AGACTACCAAGGTGAATATAAACAAGGACCAAGCAAACTCTTAACCTCTCTCAATTTGTC 1713
Qy 399 rValGlyLeuTyxValPheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPh 419
Db 1714 CGTCGGACTTATGTCTTCAAGTCACATGTTCTTAGTGAACACGCCCTTTGAGAAGGAT 1773
Qy 419 eValAsnValThrValLysProAlaArgValAsnLeuProProValAlaValSe 439
Db 1774 TGTCAATGTCACTGTTAAGCCTGCAGAAAGAGTCAACCTGCCACCTGTAGCAGTTGTTTC 1833
Qy 439 rProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuLeuAspGlySerGlnSe 459
Db 1834 TCCCAACTGCAAGAGCTCACTTTGCCCTTTGACGTCAAGCCCTCATTTGATGGACCCAAAG 1893
Qy 459 rThrAspAspThrGluIleValSerTyxHisTrpGluIleAsnGlyProPheIleG1 479
Db 1894 TACAGATGATACTGAATAGTAGTATCATTTGGGAAGAATAAAGCGGCCCTTCATAGA 1953
Qy 479 uGluLysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTy 499
Db 1954 AGAGAAGACTTTCAGTTGACTCTCCGCTTTCAGCTTGTCTAACTTGTATCTGGTAACTA 2013
Qy 499 rSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrAlaAlaLe 519
Db 2014 TAGTTTCAGGTTGACTGTTTACAGACTCGGACGGAGCCCACTAACTCTCAACTGCGAGCCCT 2073
Qy 519 uIleValAsnAsnAlaValAspTyxProProValAlaAsnAlaGlyProAsnHisThrI1 539
Db 2074 AATAGTGAACTACTGTGGACTACCCACAGTTGCTTAATCGAGGACCAATACACCAT 2133
Qy 539 eThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnI1 559
Db 2134 AACTTTGCCCAAACTCCATCACTTTGAATGGAAACCAGAGCAGTGCAGCATCACAGAT 2193
Qy 559 eValLeuTyxGluTrpSerLeuGlyProGlySerGluGlyLysHisValValMetGlnG1 579
Db 2194 TGTCTCTATCAGTGTGCTCCGCTGCTGGAGTGGAGGCAAAATGTTGGTCAATGAGGG 2253
Qy 579 yValGlnThrProTyxLeuHisLeuSerAlaMetGlnGluGlyAspTyxThrPheGlnLe 599
Db 2254 AGTACAGCGCATACCTTTATTTATCTGCAATGCAAGGAGAGATTATACATTTCACT 2313
Qy 599 uLysValThrAspSerArgGlnGlnSerThrAla***ValThrValIleValGlnPr 619
Db 2314 GAAGGTGACAGATCTTCAAGGCAACAGTCTACTGCTGTGTGATGCTGATGTTGCCAGCC 2373
Qy 619 oGluAsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProVa 639
Db 2374 TGAANAACAATAGACCTCCAGTGGCTGTGGCGGCCCTGATAAAGAGCTGATCTTCCCACT 2433
Qy 639 lGluSerAlaThrLeuAspGlySerSerSerSerAspAspHisGlyIleValPheTyxHi 659
Db 2434 GGAAGTCTACCTCCGATGGAGCAGCAGCAGCAGATGATGATGATGATGATGATGATGATGAT 2493
Qy 659 sTrpGluHisValArgGlyProSerAlaValAluMetGluAsnIleAspLysAlaIleAl 679
Db 2494 CTGGGAGCAGCTCAGAGGCCCGCTGAGTGGAGATGGAATATTTGACAAAGCAATAGC 2553
Qy 679 aThrValThrGlyLeuGlnValGlyThrTyHisPheArgLeuThrValLysAspGlnG1 699
Db 2554 CACTGTGACTGTCTCCAGTGGGACCTTACCACTTCCGTTTTCAGTGAAGAACCCAGCA 2613
Qy 699 nGluLeuSerSerThrLeuThrValAlaValLysLysGluAsnAsnSerProPr 719
Db 2614 GGGACTGAGCAGCAGTCCACCTCACTGTGGCTGTGAAGAGGAAATAATAGTCTCTCC 2673

Qy 719 oArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAs 739
Db 2674 CAGAGCCCGGCTGTGGCAGACATGTTCTTGTGCTTCCCAATAATTCATTAATTTGGA 2733
Qy 739 pGlySerArgSerThrAspAspGlnArgIleValSerTyxLeuTrpIleArgAspGlyG1 759
Db 2734 TGGTTCAGGCTCTACTGATGACCAAGAAATTTGTCTCTATCTGTGGATCCGGATGGCA 2793
Qy 759 nSerProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuTh 779
Db 2794 GAGTCCAGCAGCTGGAGATGTCACTGATGGCTCTGACCACAGTGTGGCTCTGCAGCTTAC 2853
Qy 779 rAsnLeuValGluGlyValTyxThrPheHisLeuArgValThrAspSerGlnGlyAlaSe 799
Db 2854 GAATCTGTGGAGGGGTGTACACTTTTCCACTTCCGAGTCCACGACAGTTCAGGGGGCTC 2913
Qy 799 rAspThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValG1 819
Db 2914 GGACACAGACTCCCACTGTGGAAAGTGACCCAGACCCCTAGGAAGAGTGGCTTGGTGA 2973
Qy 819 uLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuValAr 839
Db 2974 GCTCACCCTGCAGGTTGGTGTGGGAGCTGACAGAGCAGCGGAAGGACACCCCTTGTGAG 3033
Qy 839 gGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLysIleArgAl 859
Db 3034 GCAGCTGGCTGTGCTGAACGCTGGAGCTCGGACATTAAGGTCCAGAAAGATTCGGGC 3093
Qy 859 aHisSerAspLeuSerThrValIleValPheTyxValGlnSerArgProPheLysVa 879
Db 3094 CCATCTCGATCTCAGACCCGCTGATGTTGTTTATGTACAGAGCAGCGCCCTTCAAGGT 3153
Qy 879 lLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAs 899
Db 3154 TCTCAAGCTGTGAGTGGGCCGGAATCTGCACATGCGGCTCTCAAGAGGAGAGGCTGA 3213
Qy 899 pPheLeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerG1 919
Db 3214 CTCTCTGCTTTCAGGCTTTGAGGTTGTATACAGCAGGTTGCCCTTCTGAAGTGTCTGG 3273
Qy 919 yHisGlyHisCysAspProLeuThrLysArgCysIleCysSerHisLeuTrpMetGluAs 939
Db 3274 CCATGGTCACTGGACCCCTCACAAGCGCTGCAATTTGCTCTCACTTATGATGGAGAA 3333
Qy 939 nLeuIleGlnArgTyxIleTrpAspGlyGluSerAsnCysGluTrpSerIlePheTyxVa 959
Db 3334 CCTTATACAGCGTTATATCTGGGATGGAGAGCAACTGTGAGTGGAGTATATTCTATGT 3393
Qy 959 lThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCy 979
Db 3394 GACAGTGTGGCTTTTACTCTTATTTGTCTTAACAGGAGGTTTCACTTGGCTTTGCATCTG 3453
Qy 979 sCysCysLysArgGlnLysArgThrLysIleArgLysLysThrLysThrIleLeuAs 999
Db 3454 CTGCTGCAAAAGACAAAAGAGGACTAAATCAGGAAAAAACAAGTACACCATCTCGA 3513
Qy 999 pAsnMetAspGluGlnGluArgMetGluLeuArgProLysTyxGlyIleLysHisArgSe 1019
Db 3514 TAACATGGATGAACAGGAAGAAATGGAACCTGAGGCCCAATAATGATATCAAGCACCAG 3573
Qy 1019 rThrGluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspTh 1039
Db 3574 CACAGAGCAAACTCCAGCCTGATGGTATCCGAGTCTCGAGTTGAGTTCAGAGTACCAGGAC 3633
Qy 1039 rIlePheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerI1 1059
Db 3634 AATCTTCAGCGAAGAAAGATGGAGAGGGAATCCAAAGGTTTCCATGAATGTTCCAT 3693
Qy 1059 eArgAsnGlyAlaSerPheSerTyxSerLysAspArg 1072
Db 3694 CAGAAATGGAGCTTCTTCACTTATTTGCTCAAGGACAGA 3733

RESULT 4	Score:	5546.50	Matches:	1069
ADRO0859	Percent Similarity:	99.53%	Conservative:	0
ID ADRO0859 standard; DNA; 6797 BP.	Best Local Similarity:	99.53%	Mismatches:	3
XX	Query Match:	99.40%	Indels:	2
AC	DB:	13	Gaps:	1
XX	US-10-764-390-3 (1-1072) x ADRO0859 (1-6797)			
DT	04-NOV-2004 (first entry)			
XX	Human 254PID6B v.1 nucleotide sequence SEQ ID NO:270.			
DE	254PID6B; small interfering RNA; siRNA; immune response;			
XX	254PID6B-related protein; cytostatic; gene therapy; cancer; human;			
KW	254PID6B v.1; chromosome 6; gene; ds.			
KW	Homo sapiens.			
XX	WO2004067716-A2.			
XX	12-AUG-2004.			
XX	23-JAN-2004; 2004WO-US001965.			
PF	24-JAN-2003; 2003US-0442526P.			
XX	(AGEN-) AGENSYS INC.			
XX	Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;			
PI	Perez-Villar JJ, Faris M;			
XX	WPI; 2004-580991/56.			
XX	New 254PID6B siRNA composition comprising a double stranded siRNA that			
PT	corresponds to the nucleic acid ORF sequence which encodes the 254PID6B			
PT	protein or corresponds to a subsequence of the ORF, useful for detecting			
PT	and treating cancer.			
XX	Example 5; SEQ ID NO 270; 345pp; English.			
XX	The present invention describes a 254PID6B small interfering RNA (siRNA)			
CC	composition that comprises a double stranded siRNA that corresponds to			
CC	the nucleic acid open reading frame (ORF) sequence which encodes the			
CC	254PID6B protein, or corresponds to a subsequence of the ORF, where the			
CC	double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous			
CC	nucleotides in length. Also described: (1) a composition that comprises,			
CC	consists essentially of, or consists of a peptide of eight, nine, ten, or			
CC	eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID			
CC	NO:3, 5 or 7 ADRO0592, ADRO0594 or ADRO0596), or a peptide included in			
CC	any of the 42 lists of peptides, given in the specification, or a protein			
CC	that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or			
CC	identical to an entire amino acid sequence of PI; (2) a polynucleotide			
CC	that encodes the protein; (3) a composition comprising a polynucleotide			
CC	that is a full complement of the polynucleotide described above; (4)			
CC	generating a mammalian immune response directed to the protein of PI; (5)			
CC	detecting, in a sample, the presence of a 254PID6B-related protein or a			
CC	254PID6B-related polynucleotide; (6) a composition that modulates the			
CC	status of a cell that expresses a protein of PI; (7) a pharmaceutical			
CC	composition that comprises the composition described above in a human			
CC	unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)			
CC	a non-human transgenic animal that produces the antibody; (10) a			
CC	hybridoma that produces the antibody; (11) delivering a cytotoxic agent			
CC	or a diagnostic agent to a cell that expresses the protein of PI; and			
CC	(12) inhibiting growth, reproduction or survival of cancer cells that			
CC	express the protein of PI. 254PID6B has cytostatic activity, and can be			
CC	used in gene therapy. The compositions, molecules and methods are useful			
CC	for treating and detecting cancer. The present sequence represents the			
CC	human 254PID6B v.1 nucleotide sequence, which is used in the			
CC	exemplification of the present invention. The human 254PID6B gene is			
CC	located on chromosome 6p22.			
XX	Sequence 6797 BP; 1812 A; 1542 C; 1687 G; 1756 T; 0 U; 0 Other;			
SQ				
Alignment Scores:	0	Length:	6797	
Pred. No.:				

Qy	1	MetAlaProThrGlyValLeuSerSerLeuLeuLeuValThrIleAla----	G1	19
Db	514	ATGGGCCCCCAGAGGTGCTCTTCATTCCTGCTGCTGGTGCACAAATTCAGTTGG		573
Qy	19	YCysAlaArgLysGlnCysSerGluGlyArgThrTyzSerAsnAlaValIleSerProAs		39
Db	574	TTGTGCCGTAAGCAGTGCAGCGAGGGGAGACATATTCCAATTCAGTCAITTCACCTAA		633
Qy	39	nLeuGluThrThrArgIleMetArgValSerHisThrPheProValValAspCysThrAl		59
Db	634	CTTGGAACACACAGAAATCATGCGGTGTCTTCACACCTTCCCTGCTGCTAGACTGCACGCG		693
Qy	59	aAlaCysCysAspLeuSerSerCysAspLeuAlaTTPTrPheGluGlyArgCysTyLe		79
Db	694	CGTTTGCTGTGACCTGTCCAGCTGTGACCTGGCTGGCTGGTTCGAGGGCCGCTGTACTCT		753
Qy	79	uValSerCysProHisLysGluAsnCysGluProLysLysMetGlyProIleArgSerTy		99
Db	754	GGTGAGCTGCCCCCACAAAGAGAACTGTGAGCCCAAGAGATGGGCCCATCAGGTCTTA		813
Qy	99	rLeuThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMe		119
Db	814	TCTCAGCTTTTGTCTCGGCGCTGTTTCAGAGCCCTGCACAGCTGTCTGGACTATGGGGACAT		873
Qy	119	tMetLeuAsnArgGlySerProSerGlyIleTTPTrGlyAspSerProGluAspIleArgLy		139
Db	874	GATCTGAACAGGGGCTCCCCCTCGGGGATCTGGGGGACCTCAGCTGAGGATATCAGAA		933
Qy	139	sAspLeu***PheLeuGlyLysAspTrpGlyLeuGluGluMetSerGluTy***AspAs		159
Db	934	GGACTTGCCCTTTCTAGGCNAAGATTGGGGCTAGAGGAGATGCTGAGTACTCAGATGA		993
Qy	159	pTyArgGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySe		179
Db	994	CTACCGGAGCTGGAGAGAGGACCTCTTGCAACCCAGTGGCAAGCAGGAGCCAGAGGGAG		1053
Qy	179	rAlaGluThrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerVa		199
Db	1054	TGCCGAGTACAGGACTGGGGCTTACTCCGGGACGAGGGGGCTTCAACTCTCTGT		1113
Qy	199	lGlyAspSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyLeuAs		219
Db	1114	TGGAGACAGTCTCGGGTGCCAGCGGAGACGAGCAGGAGCCCTGAGCTCCATTACCTGAA		1173
Qy	219	nGluSerAlaSerThrProAlaProLysLeuProLysArgSerValLeuLeuProLeuPr		239
Db	1174	TGAGTCGGCTTCAACCCCTGCCCAAAACTCCCTGAGAGAAAGTGTGTCTTCCCTTGGC		1233
Qy	239	oThrThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGlnGlu		259
Db	1234	GACTACTCACTTCAGAGAGGTGTTCGAGNAAGAAAGGCTTTCAGCTCCAGGAACA		1293
Qy	259	nSerSerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProProAlaSe		279
Db	1294	ATCCAGCAACAGCTCTGGAAAAGAGGTTCTTAATGCTTCCCATAGTCTTCCCTCCGCAAG		1353
Qy	279	rLeuGluLeuSerSerValThrValGluLysSerProValLeuThrValThrProGlySe		299
Db	1354	CCTGGAGCTCAGCTCAGTCACCGTGAGAAAAAGCCAGTGCAGTCAACCCCGGGAG		1413
Qy	299	rThrGluHisSerIleProThrProThrSerAlaAlaProSerGluSerThrProSe		319
Db	1414	TACAGACACAGCATCCCAACACCTCCCACTAGCGAGCCCCCTCTGAGTCCACCCATC		1473
Qy	319	rgLeuLeuProfileSerProThrThrAlaProArgThrValLysGluLeuThrValSerAl		339

Db 1474 TGAGCTACCCATATCTCTTACCACTGCTCCAGGACAGTGAAGAACTTACGGTATCGGC 1533
Qy 339 aGlyAspAsnLeuIleIleThrLeuProAspAsnGluValGluLeuIysAlaPheValAl 359
Db 1534 TGGAGATAACCTAATTAACCTTTACCGCAATGAAGTTGAACTGAAGGCGCTTTGTTGC 1593
Qy 359 aProAlaProValGluThrThrTyraSntyGluTrpAsnLeuIleSerHisProTh 379
Db 1594 GCCAGCGCCACTGTAGAAACAACCTACAACTATGAATGGAATTTAATAAGCCACCCAC 1653
Qy 379 rAspTyrglnGlyGluIleIysGlnGlyHisGlnThrLeuAsnLeuSerGlnLeuSe 399
Db 1654 AGACTACCAAGTGAATTAACAAGGACACACCAACAACTCTTAACCTCTCTCAATTGTC 1713
Qy 399 rValGlyLeuTyrrValPheIysValThrValSerSerGluAsnAlaPheGlyGluGlyPh 419
Db 1714 CGTCGGACTTATGCTCTCAAGTCACTGTTCTTAGTGAAAACGCCCTTGAGAAGGATT 1773
Qy 419 eValAsnValThrValIysProAlaArgArgValAsnLeuProProValAlaValSe 439
Db 1774 TGTCAATGTCACTTTAAGCGCTGCCAGAAGAGTCAACCTGCACTGTAGCAGTTGTTTC 1833
Qy 439 rProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSe 459
Db 1834 TCCCAACTGCAGAGCTCACTTTGCCCTTGACGCTCAGCCCTCAATTGATGCGACCCAAAG 1893
Qy 459 rThrAspAspThrGluIleValSerTyrrHisTrpGluGluIleAsnGlyPropheIleGl 479
Db 1894 TACAGATGATACTGAATAGTGAGTTATCACTGGGAAGAAATAAACGGGCCCTTCATAGA 1953
Qy 479 uGluIysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTy 499
Db 1954 AGAAGAAGCTTCAGTTGACTCTCCGCTTACGCTTGTCTAAACCTTGATCCGTGTAACCTA 2013
Qy 499 rSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLe 519
Db 2014 TAGTTTCAGGTTGACTGTTACAGACTCGAGCGAGCCACTAACTCTCAACTGAGCGCCT 2073
Qy 519 uIleValAsnAsnAlaValAspTyrrProProValAlaAsnAlaGlyProAsnHisThrIl 539
Db 2074 AATAGTGAACAATGCTGTGGACTACCCACCAGTTGCTAATGCAGGACCAATCACACAT 2133
Qy 539 eThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIl 559
Db 2134 AACTTTGCCCCAAAACCTCCATCACTTTGAATGGAAACAGAGCAGTCAACATCACAGAT 2193
Qy 559 eValLeuTyrrGluTrpSerLeuGlyProGlySerGluGlyIysHisValValMetGlnI 579
Db 2194 TGTCTCTATAGTGGTCCCTCGGTCTCTGGAGTGAGGGGCAAAACATGTGGTCAATGCAAGG 2253
Qy 579 yValGlnThrProTyrrLeuHisLeuSerAlaMetGlnGluGlyAspTyrrThrPheGlnLe 599
Db 2254 AGTACAGCGCCATACCTTCATTTATCTGCAATGCAGGAGGAGATTAATACATTTACGCT 2313
Qy 599 uIysValThrAspSerArgGlnGlnSerThrAla***ValThrValIleValGlnPr 619
Db 2314 GAAGGTGACAGATTCTTCAGGCAACAGCTACTGCTGTGTGCTGCTGTGATGTTGCCAGCC 2373
Qy 619 oGluAsnAsnArgProProValAlaValAlaGlyProAspIysGluIlePheProVa 639
Db 2374 TGAATAACAATAGACTCCAGTGGCTGTGGCGGCCCTGATAAAGAGCTGATCTTCCAGT 2433
Qy 639 lGluSerAlaThrLeuAspGlySerSerSerAspAspHisGlyIleValPheTyrrHi 659
Db 2434 GGAAGTGTACCTCTGATGGAGCAGCAGCAGCGATGACCCAGCATTTGCTCTACCA 2493
Qy 659 sTrpGluHisValArgGlyProSerAlaValGluMetGluAsnIleAspIysAlaIleAl 679
Db 2494 CTGGGAGCAGCTCAGAGGCCCTCAGTGCAGTGGAGATGGAAATATTGACAAAGCAATAGC 2553
Qy 679 aThrValThrGlyLeuGlnValGlyThrTyrrHisPheArgLeuThrValIysAspGlnGl 699

Db 2554 CACTGTGACTGGTCTCCAGGTGGGACCTACCACCTTCGGTTTGACAGTGAAGACCAAGCA 2613
Qy 699 nGlyLeuSerSerThrSerThrLeuThrValAlaValIysGlnGluAsnSerProPr 719
Db 2614 GGGACTGAGCAGCAGCTCCACCCTCACTGTGGCTGTGAAGAGGAAAAATAATAGTCTCTCC 2673
Qy 719 oArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAs 739
Db 2674 CAGAGCCCGGCTGGTGGCAGACATGTTCTTGTGCTTCCCAATAATTCATTAATCTTGGGA 2733
Qy 739 pGlySerArgSerThrAspAspGlnArgIleValSerTyrrLeuTrpIleArgAspGlyGl 759
Db 2734 TGGTTCAGGCTCACTGATGATCCAAAGAAATTTGTCTCTATCTGTGGATCCGGATGGCA 2793
Qy 759 nSerProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuTh 779
Db 2794 GAGTCCAGCAGCTGGAGATGTCATGGCTCTGCACACAGATGTGGCTCTGCAGCTTAC 2853
Qy 779 rAsnLeuValGluGlyValTyrrThrPheHisLeuArgValThrAspSerGlnGlyAlaSe 799
Db 2854 GAATCTGGTGGAGGGGTGTACACTTTTCCACTTTCGAGTCAACAGTCAAGGGGGCCTC 2913
Qy 799 rAspThrAspThrAlaThrValGluValGlnProAspProArgIysSerGlyLeuValGl 819
Db 2914 GGACACAGACACTGCCACTCTGGAGTGCAGCCAGACCTTAGGAAGAGTGGCCTTGGTGA 2973
Qy 819 uLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgIysAspThrLeuValAr 839
Db 2974 GCTGACCTCGAGTTGGTGTGGCAGCTGCACAGCAGCGGAGGACACCCCTTGTGAG 3033
Qy 839 gGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleIysValGlnIysIleArgAl 859
Db 3034 GCAGCTGGCTGTGCTGCTGAACGCTGGACTCGGACATTAAGGTCCAGAAGATTCGGGC 3093
Qy 859 aHisSerAspLeuSerThrValIleValPheTyrrValGlnSerArgProPheIysVa 879
Db 3094 CCACTCGATCTCAGCACCGCTGATTTGTTTATGTACAGCAGCGCGCCTTTCAGGT 3153
Qy 879 lLeuIysAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerIysGluIysAlaAs 899
Db 3154 TCTCAAGCTGCTGAAGTGGCCGAAATCTGCACATCGGCTCTCAAAAGGAGAGGTGA 3213
Qy 899 pPheLeuLeuPheIysValLeuArgValAspThrAlaGlyCysLeuLeuIysCysSerGl 919
Db 3214 CTTCTGCTTTTCAAGGCTTGAGGGTTGATACAGCAGGTGGCCTTCTGAAGTGTCTGG 3273
Qy 919 yHisGlyHisCysAspProLeuThrIysArgCysIleCysSerHisLeuTrpMetGluAs 939
Db 3274 CCATGTCATCGACCCCTCTCAAAAGCGCTGCATTTGCTCTCCTATATGGATGGAGAA 3333
Qy 939 nLeuIleGlnArgTyrrIleTrpAspGlyGluSerAsnCysGluTrpSerIlePheTyrrVa 959
Db 3334 CCTTATACAGCGTTATATCTGGGATGAGAGAGCAACTGTGTGAGTATATTTCTATGT 3393
Qy 959 lThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCy 979
Db 3394 GACAGTGTGGCTTTACTCTTATTTGCTTAACAGGAGGTTTCACTTGGCTTTCGACTGT 3453
Qy 979 sCysCysLeuArgGlnIysArgThrIysIleAtqIysLysThrIysTyrrThrIleLeuAs 999
Db 3454 CTGCTGCAAAAGCAAAAAAGGACTAAAATCAGGAAAAAACAAGTACACCATCTCTGGA 3513
Qy 999 pAsnMetAspGluGlnGluArgMetGluLeuArgProIysTyrrGlyIleIysHisArgSe 1019
Db 3514 TAACATGATGAACAGAGAAAGATGGAACCTGAGSCCCCAATAATGTATCAAGCACCCGAG 3573
Qy 1019 rThrGluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspTh 1039
Db 3574 CACAGACACAACCTCCAGCCTGTATGTTATCCGAGTCTCGAGTTTGCAGTGCACAGGACAC 3633
Qy 1039 rIlePheSerArgGluIysMetGluArgGlyAsnProIysValSerMetAsnGlySerIl 1059
Db 3634 AATCTTCAGCCGAAAAAGATGGAGAGGGGAATCCAAAGGTTTCCATGAATGTTCCAT 3693

QY 1059 eArgAsnGlyAlaSerPheSerTyrCysSerLysAspArg 1072
 |||||
 Db 3694 CAGAATGGAGCTTCCTTCAGTTATGCTCAAGGACAGA 3733

RESULT 5

ADRO0595
 ID ADRO0595 standard; cDNA; 6991 BP.
 XX
 AC ADRO0595;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Human 254PID6B v.3 encoding cDNA SEQ ID NO:6.
 XX
 KW 254PID6B; small interfering RNA; siRNA; immune response;
 KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
 KW 254PID6B v.3; gene; ss; chromosome 6.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 739..3930
 FT /*tag= a
 FT /product= "254PID6B v.3"

XX WO2004067716-A2.

XX 12-AUG-2004.

XX 23-JAN-2004; 2004WO-US001965.

XX 24-JAN-2003; 2003US-0442526P.

XX (AGEN-) AGENSYS INC.

XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Faris M;

XX WPI; 2004-580991/56.
 DR P-PSDB; ADRO0596.

XX New 254PID6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

PS Claim 10; SEQ ID NO 6; 345pp; English.

XX The present invention describes a 254PID6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID
 CC NO:3, 5 or 7 ADRO0592, ADRO0594 or ADRO0596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of PI; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of PI; (5)
 CC detecting, in a sample, the presence of a 254PID6B-related protein or a
 CC 254PID6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of PI; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of PI; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that

CC express the protein of P1. 254PID6B has cytosolic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence encodes the human
 CC 254PID6B v.3, which is used in the exemplification of the present
 CC invention. The human 254PID6B gene is located on chromosome 6p22.
 XX

SQ Sequence 6991 BP; 1873 A; 1589 C; 1738 G; 1791 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 6991
 Score: 5527.00 Matches: 1068
 Percent Similarity: 96.83% Conservatives: 0
 Best Local Similarity: 96.83% Mismatches: 4
 Query Match: 99.05% Indels: 31
 DB: 13 Gaps: 1

US-10-764-390-3 (1-1072) x ADRO0595 (1-6991)

QY 1 MetAlaProThrGlyValLeuSerSerLeuLeuLeuValThrIleAla----- 18
 |||||
 Db 621 ATGGCGCCCCCACAGGTGTCTCTCTTCAATTGCTGCTGTGTCACAAATTCAGTTTGC 680
 |||||
 QY 18 ----- 18
 |||||
 Db 681 TTATGTTGATGTCATCATGGCAAAAAAATCACTGGTGAGCATCATTTAAGAAGACCCAT 740
 |||||
 QY 19 -----GlyCysAlaArgLysGlnCysSerGluGlyArgTh 30
 |||||
 Db 741 GACTAGACTGGCTGGCGGAGCCCATGTTGTGCGGTAAAGCAGTGCGAGGAGGAGGAC 800
 |||||
 QY 30 rTyrSerAsnAlaValIleSerProAsnLeuGluThrThrArgIleMetArgValSerHi 50
 |||||
 Db 801 ATATTCCAATGCAGTCATTTCACTTGGAAACACACAGAAATCATGCGGTGTCTCA 860
 |||||
 QY 50 sThrPheProValValAspCysThrAlaAlaCysCysAspLeuSerSerCysAspLeuAl 70
 |||||
 Db 861 CACCTTCCTCTGCTAGACTGACGGCGCTTGTGTGACCTGTCTCAGCTGTGACCTGGC 920
 |||||
 QY 70 aTrrTrrPheGluGlyArgCysTyrLeuValSerCysProHisLysGluAsnCysGluPr 90
 |||||
 Db 921 CTGGTGGTTTCAGGGCGCTCTCTACCTGGTGAGTGGCCCCCAAAAGAGAACTGTGAGCC 980
 |||||
 QY 90 oLysLysMetGlyProIleArgSerTyrLeuPheValLeuArgProValGlnArgPr 110
 |||||
 Db 981 CAAGAAGATGGCCCCCATCAGTCTTATCTCACTTTTGTCTCCGGCTGTTCAGAGGCC 1040
 |||||
 QY 110 oAlaGlnLeuLeuAspTyrGlyAspMetLeuAsnArgLysSerProSerGlyIleTr 130
 |||||
 Db 1041 TGCACAGCTGTGGAGCATATGGGAGCATATGCTGAACAGGGGCTCCCCCTCGGGGATCTG 1100
 |||||
 QY 130 pGlyAspSerProGluAspIleArgLysAspLeu***PheLeuGlyLysAspTrrpGlyLe 150
 |||||
 Db 1101 GGGGAGCTCACTGAGGATATCAGAAAGACTTGGCCCTTTCTAGGCAAGATTTGGGGCT 1160
 |||||
 QY 150 uGluGluMetSerGluTyr***AspAspTyrArgGluLeuGluLysAspLeuGlnPr 170
 |||||
 Db 1161 AGAGGAGATGCTGAGTACTCAGATGACTACCGGGAGCTGGAGAAGGACCTCTTCAACC 1220
 |||||
 QY 170 oSerGlyLysGlnGluProArgGlySerAlaGluTyrThrAspTrrpGlyLeuLeuProGl 190
 |||||
 Db 1221 CAGTGGCAAGCAGGAGCCAGAGGAGTGGCGAGTACAGGACTGGGGCTTACTGCGGG 1280
 |||||
 QY 190 ySerGluGlyAlaPheAsnSerSerValGlyAspSerProAlaValProAlaGluThrGl 210
 |||||
 Db 1281 CAGCGAGGGGCTTCAACTCTCTGTGGAGACAGTCTCTGGGTGGCAGCGGAGAGCGCA 1340
 |||||
 QY 210 nGlnAspProGluLeuHisTyrLeuAsnGluSerAlaSerThrProAlaProLysLeuPr 230
 |||||
 Db 1341 GCAGGACCTGAGCTCCATTTACTTGAATGAGTGGCTTCAACCCCTGCCCAAACTCCC 1400
 |||||
 QY 230 oGluArgSerValLeuLeuProLeuProThrThrProSerSerGlyGluValGluGly 250
 |||||
 Db 1401 TGAGAGAAAGTGTGTGCTTCCCTTCCCGACTACTCTCACTTTCAGGAGAGGTGTGAGAA 1460
 |||||

Qy 250 sGluIysAlaSerGlnLeuGlnGlnSerSerAsnSerSerGlyIysGluValLeuMe 270
Db 1461 AGAAAGAGCTCTCAGCTCCAGGAACAATCCAGCAACAGCTCTCGAAAGAGGTTCTAAT 1520
Qy 270 tProSerHisSerLeuProProAlaSerLeuGluLeuSerSerValThrValGluIysSe 290
Db 1521 GCCTTCCCATAGTCTTCTCCGGCAGCTCGAGCTCAGCTCAGTCACCGTGGGAAAG 1580
Qy 290 rProValLeuThrValThrProGlySerThrGluHisSerIleProThrProThrSe 310
Db 1581 CCCAGTGTCTACAGTCAACCCCGGAGGTACAGAGCACAGCATCCCAACACCTCCCACTAG 1640
Qy 310 rAlaAlaProSerGluSerThrProSerGluLeuProIleSerProThrThrAlaProAr 330
Db 1641 CGCAGCCCTCTCAGTCCACCCCTACCTCTGAGCTACCCATATCTCCTACCACTGTCTCCAG 1700
Qy 330 gThrValIysGluLeuThrValSerAlaGlyAsnLeuIleThrLeuProAspAs 350
Db 1701 GACAGTGAAGAACTTACGGTATCGCTGGAGATAACCTAATTAATTAATTTACCCGACAA 1760
Qy 350 nGluValGluLeuIysAlaPheValAlaProAlaProValGluThrThrTrpAsnTy 370
Db 1761 TGAAGTTCAACTGAAGGCTTTGTTGGCCAGGCGCACCTGTAGAAAACAACCTACAACCTA 1820
Qy 370 rGluTrpAsnLeuIleSerHisProThrAspTyrGlnGlyGluIleIysGlnGlyHisIy 390
Db 1821 TGAATGGAAATTAATAAGCCACCCACAGACTTACCAAGGTGAATAAACAAGGACACAA 1880
Qy 390 sGlnThrLeuAsnLeuSerGlnLeuSerValGlyLeuTyThrValPheIysValThrValSe 410
Db 1881 GCAAACTCTTAACCTCTCAATTTGTCGTGGACTTTATGTCTTCAANGTCACTGTTTC 1940
Qy 410 rSerGluAsnAlaPheGlyGluGlyPheValAsnValThrValIysProAlaArgVa 430
Db 1941 TAGTGAAGAAAGCTTTGGAGAGGATTTGTCAATGTCTACTGTTAAGCTGCCAGAGAGT 2000
Qy 430 lAsnLeuProProValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuth 450
Db 2001 CAACCTGCACCTGTAGCAGTGTGTTCTCCCAACTGCAAGAGCTCACTTTGCCCTTTGAC 2060
Qy 450 rSerAlaLeuIleAspGlySerGlnSerThrAspAspThrGluIleValSerTyHisTr 470
Db 2061 GTCAGCCCTCAATGATGCGACGCAAGATGACAGATGATCACTGAATAGTATCATTTG 2120
Qy 470 pGluGluIleAsnGlyProPheIleGluGluIysThrSerValAspSerProValLeuAr 490
Db 2121 GGAAGAAATAACGGGCCCTTCATAGAGAGAGAGACTTCAGTTGACTCTCCCGTCTTACG 2180
Qy 490 gLeuSerAsnLeuAspProGlyAsnTySerPheArgLeuThrValThrAspSerAspGl 510
Db 2181 CTTGTCTAACTTGATCTGTGTAATACTATAGTTTTCAGGTTGACTGTTACAGACTCGGACGG 2240
Qy 510 yAlaThrAsnSerThrAlaAlaLeuIleValAsnAsnAlaValAspTyProProVa 530
Db 2241 AGCCACTAACTCTCAACTGACGCGCTTAATAGTGAACAATGCTGTGGACTACCCACCACT 2300
Qy 530 lAlaAsnAlaGlyProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGl 550
Db 2301 TGCTAATCAGGACCAATACACATTAACCTTTGCCCAAACTCCATCACTTTGAATGG 2360
Qy 550 yAsnGlnSerSerAspAspHisGlnIleValLeuTyTrpSerLeuGlyProGlySe 570
Db 2361 AAACAGAGCAGTGACGATCACAGATTGCTCTATGAGTGGTCCCTGGGTCTCTGGAG 2420
Qy 570 rGluGlyIysHisValValMetGlnGlyValGlnThrProTyThrLeuHisLeuSerAlaMe 590
Db 2421 TGAGGGCAAAATGTGGTCATGCGAGGAGTACAGACGCCATACCTTCATTATCTGCAAT 2480
Qy 590 tGlnGluGlyAspTyThrPheGlnLeuIysValThrAspSerSerArgGlnGlnSerTh 610
Db 2481 CGAGGAGAGATTATATCATTTTCACTGAAGGTGACAGATTCTTCAAGGCAACAGTCTAC 2540

Qy 610 zAla***ValThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaGl 630
Db 2541 TGCTGTGTGACTGTGATTTGTCCAGCCCTGAAAACAATAGACCTCAGTGGCTGTGGCCGG 2600
Qy 630 yProAspIysGluLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSe 650
Db 2601 CCCTGATTAAGAGCTGATCTTCCAGTGGAAAGTGTACTCCCTGGATGGGAGCAGCAGAG 2660
Qy 650 rAspAspHisGlyIleValPheTyHisTrpGluHisValArgGlyProSerAlaValGl 670
Db 2661 CGATGACCAACGGCATTTCTTACCACTGGGAGCAGCTCAGAGGCCCCAGTCAGAGTGA 2720
Qy 670 uMetGluAsnIleAspIysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyHi 690
Db 2721 GATGGAATAATTTGACAAAGCAATAGCCACTGTGACTGGTCTCCAGGTTGGGACCTACCA 2780
Qy 690 sPheArgLeuThrValIysAspGlnGlnGlyLeuSerSerThrThrLeuThrValAl 710
Db 2781 CTTCCGTTTGACAGTGAAGACCAAGCAGGAGCTGAGCAGCAGCTCCACCTCAGTGTGC 2840
Qy 710 aValIysGluAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuVa 730
Db 2841 TGTGAAGAGGAAATAATAGTCTCCAGAGCCCGGCTGGTGGCAGACATGTTCTTGT 2900
Qy 730 lLeuProAsnAsnSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleVa 750
Db 2901 GCTTCCCAATAATTTCCATTTACTTTGGATGGTTCAGGTTCTACTGATGACCAAGAAATGT 2960
Qy 750 lSerTyLeuTrpIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySe 770
Db 2961 GTCTATCTGTGGATCCGGATGGCCAGAGTCCAGAGCTGAGAGTGTATCATCGATGCTC 3020
Qy 770 rAspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyValTyThrPheHisIe 790
Db 3021 TGACCACAGTGTGCTCTGCAGCTTACGAATCTGGTGGAGGGGTGTACACTTTCACCTT 3080
Qy 790 uArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnPr 810
Db 3081 GCGAGTCAACGACAGTCAAGGGGCTCGGACACAGACACTGCCACTGTGGAAAGTGCAGCC 3140
Qy 810 oAspProArgIysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuTh 830
Db 3141 AGACCTTAGNAGAGTGGCTGTGGAGCTGACCTGCAGGTTGGTGTGGCAGCTGAC 3200
Qy 830 rGluGlnArgIysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSe 850
Db 3201 AGACAGCGGAAGACACCTTGTGAGCAGCTGGCTGTGCTGTGTAACGCTGTGACTC 3260
Qy 850 rAspIleIysValGlnIysIleArgAlaHisSerAspLeuSerThrValIleValPheTy 870
Db 3261 GGACATTAAGGTCCAGAGATTCGGGCCCACTCGGATCTCAGCACCGGTGATTGTGTTTA 3320
Qy 870 rValGlnSerArgProPheIysValLeuIysAlaAlaGluValAlaArgAsnLeuHi 890
Db 3321 TGTCAGACAGCGCGCTTTCAAGGTTCTCAAGCTGTGTGAGTGGCCCAATCTGCA 3380
Qy 890 sMetArgLeuSerIysGluIysAlaAspPheLeuLeuPheIysValLeuArgValAspTh 910
Db 3381 CATCGGCTCTCAAGAGAGAGGCTGACTTCTGCTTTTCAAGTCTTTAGGGTTGATAC 3440
Qy 910 rAlaGlyCysLeuLeuIysCysSerGlyHisGlyHisCysAspProLeuThrIysArgCy 930
Db 3441 AGCAGGTTGCCCTTCTGAAGTGTCTGGCCATGGTCACTCGCACCCCTTCACAAAGCGCTG 3500
Qy 930 sIleCysSerHisLeuTrpMetGluAsnLeuIleGlnArgTyTrpAspGlyVGluse 950
Db 3501 CATTGTCTCTCATTTAGATGGAGAACTTATACAGGTTATATCTGGGATGGAGAGAG 3560
Qy 950 rAsnCysGluTrpSerIlePheTyThrValThrValLeuAlaPheThrLeuIleValLeuTh 970
Db 3561 CAACGTGAGTGGAGTATATCTATGTGACAGTGTGGCTTTTACTTTATTTGTGCTAAC 3620
Qy 970 rGlyGlyPheThrTrpLeuCysIleCysCysArgGlnIysArgThrIysIleAr 990

Db 1221 CAGTGGCAAGCAGGAGCCAGAGGAGTGGCCGAGTACAGGACTGGGCGCTACTGGCCGG 1280
Qy 190 ySerGluGlyAlaPheAenSerSerValGlyApsSerProAlaValProAlaGluThrG1 210
Db 1281 CAGGAGGGGGCCCTTCACTCCTCTGTTGGAGACAGTCTCGCGGTGGCCAGCGAGCGCA 1340
Qy 210 nGlnApsProGluLeuHisTyrLeuAenGluSerAlaSerThrProAlaProLysLeuPr 230
Db 1341 CGAGGACCTGAGCTCCATTAACCTGMAATGAGTGGCTTCAACCCCTGCCCAAACTCCC 1400
Qy 230 oGluApsSerValLeuLeuProLeuProThrThrProSerSerGlyGluValLeuGluLy 250
Db 1401 TGAGAGAAGTGTGTGCTTCTCCCTGGCAGTACTCCATCTTTCAGGAGAGGTGTTGGAGAA 1460
Qy 250 sGluLysAlaSerGlnLeuGlnSerSerAenSerSerGlyLysGluValLeuMe 270
Db 1461 AGAAAGGCTTCTCAGCTCCAGGAAACAATCCAGCAACAGCTCTCGAAAGAGGTCTTAAT 1520
Qy 270 tProSerHisSerLeuProAlaSerLeuGluLeuSerSerValThrValGluLysSe 290
Db 1521 GCCTTCCCATAGTCTTCTCCGGCAGCCTGGAGCTCAGCTCAGTCCAGCTGGGAGAAAG 1580
Qy 290 rProValLeuThrValThrProGlySerThrGluHisSerIleProThrProThrSe 310
Db 1581 CCCAGTGTCTACAGTCAACCCCGGGAGTACAGAGCACAGCATCCCAACACCTCCCACTAG 1640
Qy 310 rAlaAlaProSerGluSerThrProSerGluLeuProIleserProThrThrAlaProAr 330
Db 1641 CGCAGCCCCCTCTCAGTCCACCCATCTGAGCTACCCATATCTCTCACCACCTGCTCCAG 1700
Qy 330 qThrValLysGluLeuThrValSerAlaGlyApsAenLeuIleIleThrLeuProApsAs 350
Db 1701 GACAGTGAAGAACCTTACGGTATCGGCTGGAGATAACCTTAATTAACCTTTACCCGCAAA 1760
Qy 350 nGluValGluLeuLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTy 370
Db 1761 TGAAGTTGAACCTGAAGGCTTGTGTGGCCAGCGCCACCTGTAGAAACAACCTACACTA 1820
Qy 370 rGluTrpAsnLeuIleserHisProThrApsTyrGlnGlyGluIleLysGlnGlyHisLy 390
Db 1821 TGAATGGAAATTAATAGCCACCCACAGACTACCAAGGTGAATTAATAAACAAGACACAA 1880
Qy 390 sGlnThrLeuAenLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSe 410
Db 1881 GCAAACTCTTAACCTCTCTCAATTTGTCGGTCGAGCTTATATCTCTCAAGTCACTGTTTC 1940
Qy 410 rSerGluAAsnAlaPheGlyGluGlyPheValAasnValThrValLysProAlaArgVa 430
Db 1941 TAGTGAANAACCCCTTTGGAGAGAGATTTGTCAATGTCACTGTTAAGCTGCCAGAGAGT 2000
Qy 430 lAasnLeuProProValAlaValSerProGlnLeuGlnGluLeuThrLeuProLeuTh 450
Db 2001 CAACCTGCCACCTGTAGCAGTTGTTCTTCCCAACTGCAAGAGCTCACTTTGCGCTTTGAC 2060
Qy 450 rSerAlaLeuIleApsGlySerGlnSerThrApsApsThrGluIleValSerTyrHisTr 470
Db 2061 GTACGCCCTCATGTAGCCACCCAAAGTACAGATGATACGAAATAGTGAATTTATCATTTG 2120
Qy 470 pGluGluIleAasnGlyProPheIleGluGluLysThrSerValApsSerProValLeuAr 490
Db 2121 GGAAGAAATAAACCGGCCCTTCATAGAGAGAGACTTCAGTTGACTGTTCAGACTCCCGCTTACG 2180
Qy 490 gLeuSerAasnLeuApsProGlyAasnTyrSerPheArgLeuThrValThrApsSerApsG1 510
Db 2181 CTTGTCTAACCTTGATCTCTGTTAATACTAGTTTTCAGGTTGACTGTTCAGACTCGGACGG 2240
Qy 510 yAlaThrAasnSerThrThrAlaAlaLeuIleValAasnAasnAlaValApsTyrProProVa 530
Db 2241 AGCCACTAACTCTCAACTGACGCCCTTAATAGTGAACAATGCTGTGGACTACCCACCAGT 2300
Qy 530 lAlaAasnAlaGlyProAenHisThrIleThrLeuProGlnAasnSerIleThrLeuAenG1 550
Db 2301 TGCTAATGCGAGCAAAATCAACATNACTTTTGGCCCAAAACTCCATCACTTTTGAATGG 2360

Qy 550 yAenGlnSerSerApsApsHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySe 570
Db 2361 AAACCAAGAGCAGTGACGATCACCAGATTGTCTCTATGAGTGGTCCCTGGGTCTCTGGAG 2420
Qy 570 xGluGlyLysHisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMe 590
Db 2421 TGAGGGCAACAATGTGTGTCATGAGGAGGTACAGCGCCATACCTTCATTTATCTGCANT 2480
Qy 590 tGlnGluGlyApsTyrThrPheGlnLeuLysValThrApsSerSerArgGlnGlnSerTh 610
Db 2481 GCAGAAAGGAGATTATACATTTCACTGAGCTGAAGTGCACAGATTTCTTCAAGGCAACAGTCTAC 2540
Qy 610 rAlaA***ValThrValIleValGlnProGluAasnAasnArgProProValAlaValAlaG1 630
Db 2541 TGCTGTGGTGAAGTGTGATTTGCCAGCTGAAACAANATAGACCTCCAGTGGCTGTGGCCGG 2600
Qy 630 yProApsLysGluLeuIlePheProValGluSerAlaThrLeuApsGlySerSerSe 650
Db 2601 CCCTGATTAAGAGCTGATCTTCCAGTGGAAAGTGTCTACCTCGATGGGAGCAGCAGNG 2660
Qy 650 rApsApsHisGlyIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValG1 670
Db 2661 CGATGACACCGCATTTGCTTCTACCACTGGGAGCAGTCCAGAGGCCCCAGTGCAGTGA 2720
Qy 670 uMetGluAasnIleApsLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHi 690
Db 2721 GATGAAATAATTATGACAAAGCAATAGCCACTGTGACTGGTCTCCAGGTGGGAGCCTACCA 2780
Qy 690 sPheArgLeuThrValLysApsGlnGlnGlyLeuSerSerThrSerThrLeuThrValAl 710
Db 2781 CTTCCGTTTGACAGTGAAGACACGAGGAGTGTGAGCAGCAGCTCCACCTCACTGTGGC 2840
Qy 710 aValLysGlyGluAasnAasnSerProProArgAlaArgAlaGlyArgHisValLeuVa 730
Db 2841 TGTGAAGAAGAAATAATAGTCTCTCCAGAGCCCGGCTGGTGGCAGACATGTTCTTGT 2900
Qy 730 lLeuProAasnAasnSerIleThrLeuApsGlySerArgSerThrApsApsGlnArgIleVa 750
Db 2901 GCTTCCCAATAATTCCATTTACTTTGGATGGTTCAAGGTCTACTGATGACCAAGAAATGT 2960
Qy 750 lSerTyrLeuTrpIleArgApsGlyGlnSerProAlaAlaGlyApsValIleApsGlySe 770
Db 2961 GTCTATCTGTGGATCCGGATGGCCAGAGTCCAGAGCTGGAGATGTCACTGATGGCTC 3020
Qy 770 rApsHisSerValAlaLeuGlnLeuThrAasnLeuValGluGlyValTyrThrPheHisLe 790
Db 3021 TGACCACAGTGTGGCTCTGCAGCTTACGAATCTGTGGAGGGGGTGTACACTTTCCACTT 3080
Qy 790 uArgValThrApsSerGlnGlyAlaSerApsThrApsThrAlaThrValGluValGlnPr 810
Db 3081 GCGAGTCACCGACAGTCAAGGGGCTCGGACACAGACACTGCCACTGTGGAAAGTGCAGCC 3140
Qy 810 oApsProArgLysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuTh 830
Db 3141 AGACCTTAGAAGAGTGGCTGTGGAGCTGACCTCGAGGTGGTGTGGGAGCTGAC 3200
Qy 830 rGluGlnArgLysApsThrLeuValArgGlnLeuAlaValLeuLeuAasnValLeuApsSe 850
Db 3201 AGAGCAGCGGAGACACCTTGTGAGGAGCTGGCTGTGCTGTGAAACGTCTGGACTC 3260
Qy 850 rApsIleLysValGlnLysIleArgAlaHisSerApsLeuSerThrValIleValPheTy 870
Db 3261 GGACATTAAGTCTCAGAAGATTGGGGCCACTCGGATCTCAGACACCGTGTGTTGTTT 3320
Qy 870 rValGlnSerArgProProPheLysValLeuLysAlaGluValAlaArgAasnLeuHi 890
Db 3321 TGTACAGAGCAGCGCCCTTTCAAGGTTCTCAAAGCTGTCTCAAAGTGGCCGAAATCTGCA 3380
Qy 890 sMetArgLeuSerLysGluLysAlaApsPheLeuLeuPheLysValLeuArgValApsTh 910
Db 3381 CATGCGGCTCTCAAGAGAGAGGCTGACTTCTTCTTCTTCTTCAAGGCTCTTGAGGGTTGATAC 3440

101 ThrPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMet 120
120 Db ACTTTTGTGCTCCGGCCCTTTTCAGAGGCTTCACAGCTGCTGGACTATGGGACATGATG 426
121 LeuAsnArgGlySerProSerGlyIleTyrGlyAspSerProGluAspIleArgLysAsp 140
120 Db CTGAACAGGGGCTCCCTCGGGATCTGGGGGACTCACCTGAGGATATCAGAAAGGAC 486
141 Leu***PheLeuGlyLysAspTyrGlyLeuGluMetSerGluTyr***AspAspTyr 160
120 Db TTGCCCTTTCTAGGCAAGATTTGGGGCTAGAGGAGATGTTCTGAGTACTCAGATGACTAC 546
161 ArgGluLeuGlyLysAspLeuGlnProSerGlyLysGlnGluProArgGlySerAla 180
120 Db CGGGAGCTGGAGAGGAGCTCTTGCACCCAGTGGCAGAGGAGGAGGAGGAGGAGTCC 606
181 GluTyrThrAspTyrGlyLeuLeuProGlySerGlyLysAlaPheAsnSerSerValGly 200
120 Db GAGTACACGAGTGGGGCTACTCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 666
201 AspSerProAlaValProAlaGluThrGlnGlnAspProGluLeuLeuHisTyrLeuAsnGlu 220
120 Db GACAGCTCTCGGGTGGCAGGAGAGCGAGCAGCAGGAGCAGCAGGAGCAGCAGGAGGAG 726
221 SerAlaSerThrProAlaProLysLeuProGluArgSerValLeuLeuProLeuProThr 240
120 Db TCGGCTTCAACCCCTCGCCCAAACTCCCTCAGAGAGTGTGTGCTTCCCTTCCCGACT 786
241 ThrProSerSerGlyGluValLeuGlyLysGlyLysAlaSerGlnLeuGlnGlnSer 260
120 Db ACTCACTTTCAGAGAGGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
261 SerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProProAlaSerLeu 280
120 Db AGCAACAGCTCTGGAAAGAGGTCTTAATGCTTCCCATAGTCTTCCCTCCCGCAGCCTG 906
281 GluLeuSerSerValThrValGlyLysSerProValLeuThrValThrProGlySerThr 300
120 Db GAGTCACTAGTCAGTCACCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966
301 GluHisSerIleProThrProThrSerAlaAlaProSerGluSerThrProSerGlu 320
120 Db GAGCAGACATCCCAACACCTCCCACTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1026
321 LeuProIleSerProThrThrAlaProArgThrValLysGluLeuThrValSerAlaGly 340
120 Db CTACCCATATCTCTACCACTGCTCCCGAGCAGTGAAGAGAGTACGGTATCGCTGGA 1086
341 AspAsnLeuIleIleThrLeuProAspAsnGluValGluLeuLysAlaPheValAlaPro 360
120 Db GATAACCTTAATATAACTTTTACCCGACAAATGAAGTGAAGTGAAGGAGGAGGAGGAGGAG 1146
361 AlaProProValGluThrThrTyrAsnTyrGluTyrAsnLeuIleSerHisProThrAsp 380
120 Db GCGCCACCTGTAGAAACCACTTACAACTATGAATGAATTAATGAATGAATTAATGAATGAAT 1206
381 TyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerVal 400
120 Db TACCAAGGTGAATATAAACAAGGACACCAAGCAAACTCTTAACCTCTCTCAATTTGCTGTC 1266
401 GlyLeuTyrValPheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPheVal 420
120 Db GGACTTTATGCTTCAAAAGTCACTGTTTCTAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1326
421 AsnValThrValLysProAlaArgValAsnLeuProProValAlaValValSerPro 440
120 Db AATGTCAGTGAAGGCTGCGCAGAGAGTCAACCTGCGCAGTGAAGGAGGAGGAGGAGGAGGAGGAG 1386
441 GlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThr 460
120 Db CAATGCAAGAGCTCACTTTGCTTTGAGCTGAGCCCTCATTTGAGTGGAGGAGGAGGAGGAGGAGGAG 1446
461 AspAspThrGluIleValSerTyrHisTrpGluGluIleAsnGlyProPheIleGluGlu 480

1447 GATGATCTGAATAGTAGTATCATTTGGAGAGAAATAACCGGCCCTTCATAGAGAG 1506
481 LysThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSer 500
1507 AAGACTTCAGTTGACTCTCCGCTCTTAACTTGTAACTTGTATCTGTGTAACCTATAGT 1566
501 PheArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuIle 520
1567 TTCAGGTGTGATGTTTTCAGACTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1626
521 ValAsnAsnAlaValAspTyrProProValAlaAsnAlaGlyProAsnHisThrIleThr 540
1627 GTGACCAATGCTGTGAGTACCCACAGTGTCTAATGCAGGAGGAGGAGGAGGAGGAGGAGGAG 1686
541 LeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleVal 560
1687 TTGCCCCAAAACCTCATCTACTTTGAATGGAAACAGAGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1746
561 LeuTyrGluTyrSerLeuGlyProGlySerGluGlyLysHisValMetGlnGlyVal 580
1747 CTCTATCAGTGTGCTTGGGTCCTGGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1806
581 GlnThrProTyrLeuHisLeuSerAlaMetGlnGlyAspTyrThrPheGlnLeuLys 600
1807 CAGAGCGCATACCTTCATTTATCTGCAATGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1866
601 ValThrAspSerSerArgGlnSerThrAla***ValThrValIleValGlnProGlu 620
1867 GTGACAGATCTTCAAGGCAACAGTCTACTGCTGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1926
621 AsnAsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProValGlu 640
1927 AACAAATAGACCTCAGTGGCTGTGGCGCCCTCATATAAGAGCTGATCTTCCCAAGTGAA 1986
641 SerAlaThrLeuAspGlySerSerSerAspAspHisGlyIleValPheTyrHisTrp 660
1987 AGTCTTACCTTGGATGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2046
661 GluHisValArgGlyProSerAlaValGluMetGluAsnIleAspLysAlaIleAlaThr 680
2047 GAGCAGCTGAGAGGCCCTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 2106
681 ValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrValLysAspGlnGly 700
2107 GTGACTGTCTCCAGTGGGAGCCTACCACTTCCGTTTGCAGTGAAGAGGAGGAGGAGGAGGAGGAGGAG 2166
701 LeuSerSerThrSerThrLeuThrValAlaValLysLysGluAsnAsnSerProProArg 720
2167 CTGAGCAGCAGCTCCCTCCTCAGTGTGCTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2226
721 AlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAsnSerIleThrLeuAspGly 740
2227 GCGCGGCTGTGGCAGCAGATGTTCTGTGCTTCCCAATAATTCATTTACCTTTGGATGCT 2286
741 SerArgSerThrAspAspGlnArgIleValSerTyrLeuTrpIleArgAspGlyGlnSer 760
2287 TCAAGGTCTACTGATGACCAAAAGATTTGTCTCTCTGTGATCCGCGATCCGCGATCCGCGATCCGCGAT 2346
761 ProAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsn 780
2347 CCAGCAGCTGAGATGTCATCGATGGCTCTGACCACTGAGTGTGGCTCTGCAGCTTACGAAT 2406
781 LeuValGluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAsp 800
2407 CTGTTGGAGGGGTGTACATTTCCACTTGGAGTCCACGAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAG 2466
801 ThrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValGluLeu 820
2467 ACAGACATGCTGCTGGAAGTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2526
821 ThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuValArgGln 840

Db 2527 ACCCTGCAGTTGGTGTGGCAGCTGACAGCAGCGAAGGACACACCCCTGTGTGAGGCG 2586
 Qy 841 LeuAlaValLeuLeuAsnValLeuAspSerApeIleLysValGlnLysIleArgAlaHis 860
 Db 2587 CTGGCTGTCTGCTGCTGAAGCTGTGAGCTGGACATTAAGGTCAGAAATTCGGGCCAC 2646
 Qy 861 SerAspLeuSerThrValIleValPheTyrValGlnSerArgProPheLysValLeu 880
 Db 2647 TCGGATCTCAGCACCGTGATGTTGTTTATGTACAGAGCAGCGCCCTTTCAGAGTTCTC 2706
 Qy 881 LysAlaIaGluValAlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAspPhe 900
 Db 2707 AAAGCTGTGAGTGGCCGGAATCTGCACATGCGGCTCTCAAGGAGGAAGGCTGACTTC 2766
 Qy 901 LeuLeuPheLysValLeuArgValAspThrAlaGlyCysLeuLeuLysCysSerGlyHis 920
 Db 2767 TTGCTTTTCAAGGCTTGAAGGTTGATACAGCAGGTGCTTCTGAAGTGTTCGGCCAT 2826
 Qy 921 GlyHisCysAspProLeuThrLysArgCysIleCysSerHisLeuTrpMetGluAsnLeu 940
 Db 2827 GGTCACTGGACCCCTCACAAGCGCTGCATTTGCTCTCACTTATGATGGAGAACCTT 2886
 Qy 941 IleGlnArgTyrIleTrpAspGlyGluSerAsnCysGluTrpSerIlePheTyrValThr 960
 Db 2887 ATACAGCGTTATATCTGGGATGGAGAGCAACTGTGA----- 2924
 Qy 961 ValLeuAlaPheThrLeuIleValLeuThrGlyGlyPheThrTrpLeuCysIleCysCys 980
 Db 2924 ----- 2924
 Qy 981 CysLysArgGlnLysArgThrLysIleArgLysLysThrLysTyrThrIleLeuAspAsn 1000
 Db 2925 -----CAAAAAGGACTAAATCAGAAAAAACAAGTACACCATCTCGGATAAC 2975
 Qy 1001 MetAspGluGlnGluArgMetGluLeuArgProLysTyrGlyIleLysHisArgSerThr 1020
 Db 2976 ATGGATGAACAGAAAGATGGAAGTGAAGTGAAGGAGGAGGATCCAAAGGTTTCCATGATGTTCCATCAGA 3035
 Qy 1021 GluHisAsnSerSerLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIle 1040
 Db 3036 GAGCAAACTCCAGCGCTGATGGTATCCGAGTCTGAGTTCGAGTGACGAGGACCAATC 3095
 Qy 1041 PheSerArgGluLysMetGluArgGlyAsnProLysValSerMetAsnGlySerIleArg 1060
 Db 3096 TTCAGCCGAGAAAGATGAGAGAGGAGGATCCAAAGGTTTCCATGATGTTCCATCAGA 3155
 Qy 1061 AsnGlyAlaSerPheSerTyrCysSerLysAspArg 1072
 Db 3156 AATGGAGCTTCCTTCAGTTATTGCTCAAGAGCAGA 3191
 RESULT 8
 ADM36226
 ID ADM36226 standard; cDNA; 3147 BP.
 XX
 AC
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Murine SST20-6 cDNA related to adipocyte function SeqID 21.
 XX
 KW gene; ss; mouse; murine; differentiation; metabolic function; adipocyte;
 KW obesity; hypertension; hyperlipaemia; diabetes; arteriosclerosis;
 KW anorectic; antidiabetic; antiarteriosclerotic; antilipaemic; hypotensive.
 XX
 OS Mus musculus.
 XX
 PN WO2004007711-A1.
 XX
 XX 22-JAN-2004.
 PD
 XX 09-JUL-2003; 2003WO-JP008690.
 XX
 XX 10-JUL-2002; 2002JP-00201856.
 PR

XX (TAKE) TAKEDA CHEM IND LTD.
 PA Matsuzawa Y, Funahashi T, Shimomura C, Furuyama N;
 PI WPI; 2004-122943/12.
 XX P-PSDB; ADM36227.
 DR
 PT Mouse membrane and secretory proteins of adipocyte origin and
 PT polynucleotides encoding them for screening compounds as remedies for
 PT obesity, diabetes, arteriosclerosis, hypertension and hyperlipemia.
 XX
 PS Disclosure; SEQ ID NO 21; 195pp; Japanese.
 XX
 CC This invention relates to novel membrane proteins that are associated
 CC with differentiation and/ or metabolic function of adipocytes, in
 CC particular of mouse origin. Specifically, it refers to the isolated
 CC nucleic acid molecules encoding all or part of these proteins,
 CC appropriate antibodies and screening assays useful for the development of
 CC drug compositions derived thereof. The present invention describes these
 CC compositions as useful for the treatment of diseases associated with
 CC abnormalities of adipocyte function, such that they can be used to
 CC prevent, treat or diagnose obesity, hypertension, hyperlipaemia, diabetes
 CC and arteriosclerosis. Accordingly, they exhibit anorectic, antidiabetic,
 CC antiarteriosclerotic, antilipaemic and hypotensive activities. This
 CC polynucleotide is a murine cDNA sequence encoding a membrane protein of
 CC the invention.
 XX
 SQ Sequence 3147 BP; 816 A; 867 C; 843 G; 621 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.85e-138 Length: 3147
 Score: 2435.50 Matches: 528
 Percent Similarity: 61.87% Conservative: 139
 Best local Similarity: 48.98% Mismatches: 303
 Query Match: 43.65% Indels: 109
 DB: 12 Gaps: 19
 US-10-764-390-3 (1-1072) x ADM36226 (1-3147)
 Qy 10 SerLeuLeuLeuValThrIleAlaGlyCysAlaArgLysGlnCysSerGluGlyArg 29
 Db 121 AGCGTTCTGTGGTGTTCACAGATGCTGATGAGAGCAGA-----TGCACACAGGGGAAG 174
 Qy 30 ThrTyrSerAsnAlaValIleSerProAsnLeuGluThrThrArgIleMetArgValSer 49
 Db 175 ACACTTTATGAGCTGGCTTGAGAACTGAGGAGAA---AATCAGCTCGGGCTTCTTGA 231
 Qy 50 HisThrPheProValValAspCysThrAlaAlaCysCysAspLeuSerCysAspLeu 69
 Db 232 GGAAGCCTGCCTTTCCACGCTGTGCGGCTGCTGCGGAGCTCTGCTGCCACGCT 291
 Qy 70 AlaTrpTrpPheGluGlyArgCysTyrLeuValSerCysProHisLysGluAsnCysGlu 89
 Db 292 CTATGCTGGCTGGAAGGATGCTTTCAGGCTGACTGCAGCAAGCCCGCAGAGCTGCCAG 351
 Qy 90 ProLysLysMetGlyProIleArgSerTyrLeuThrPheValLeuArgProValGlnArg 109
 Db 352 CTTTTTAGCAGACTCTTCCATTCCATGCTG----- 384
 Qy 110 ProAlaGlnLeuLeuAspTyrGlyAspMetMetLeuAsnArgGlySerProSerGlyIle 129
 Db 385 -----ATCAATTTTCAAAAAATCCAAACTACAGATGAT 417
 Qy 130 TrpGlyAspSerProGluAspIleArgLysAspLeu***PheLeuGlyLysAspTrpGly 149
 Db 418 TTGGGCTTCTGCTTGAAGATGATGAACACATCTTCTGAGGCTAGGC-----TGGGGC 471
 Qy 150 LeuGluGluMetSerGluTyr***AspAspTyrArgGluLeuGluLysAspLeuGln 169
 Db 472 AG-GACATCGTGGAGAG-----GCAGAG 494
 Qy 170 ProSer---GlyLysGlnGluPro-----Arg 177

[illegible]

Db	1440	CTCTGACGGGGCTACCAACTCCACCACTCGCAAGCCCTGACTGTGTGAACAAAGCTGTGGACTA	1499
Qy	527	rProProValAlaAenAlaGlyProAenHisThrIleThrLeuProGlnAenSerIleTh	547
Db	1500	CCCTCCGCTGGCCNAATGACAGCCCCCAACCAACCAAGTGATCACCCCTGCCTCAGACTCCATCAC	1559
Qy	547	rLeuAenGlyAenGlnSerSerAspAspHisGlnIleValLeuTyrGluTrpSerLeuGl	567
Db	1560	CCTCTTTGGAAACAGACGACGATACCAACGCGATCACACGACTATAGTGGTTCGTCTAG	1619
Qy	567	yProGlySerGluGlyIleHisValValMetGlnGlyValGlnThrProTyrIleuHisLe	587
Db	1620	CCCAGACAGCAAGGAGAGTGTGTGAGATGCAAGGAGTTAGAACCCGACCCCTGCAGCT	1679
Qy	587	uSerAlaMetGlnGluGlyAspTyrThrPheGlnLeuLysValThrAspSerSerArgGl	607
Db	1680	GTCCGCAATGCAAGAGGAGACTATACCTTACAGCTCACAGCTGACTGACACCGCAGGACA	1739
Qy	607	nGlnSerThrAla**ValThrValIleValGlnProGluAenAenArgProProValAl	627
Db	1740	ACAGGCCACCGCCCAAGTGACTGTGATGTGTGCAGCCTGTGAAACAAACAGCCTCTCAGGC	1799
Qy	627	aValAlaGlyProAspLysGluLeuIlePheProValGluSerAlaThrLeuAspGlySe	647
Db	1800	AGATGCAGGCCACAAAGAGCTGACCTTGCCTGCGAGCAGCAACCTGACGCGCAG	1859
Qy	647	rSerSerSerAspAspHisGlyIleValPheTyrHisTrpGluHisValArgGlyProSe	667
Db	1860	CAAGAGCACAGATGACCAGAGATCGTCTCTTACCTTTGGGACAGAGTCGGGAGCCTGA	1919
Qy	667	rAlaValGluMetGluAenIleAspLysAlaIleAlaThrValThrGlyLeuGlnValGl	687
Db	1920	CGGGGTGCAGCTGGAGAAATCCCAACACAGACTGTGCGCACTGTGACTGTGGGTGCAAGTCGG	1979
Qy	687	yThrTyrHisPheArgLeuThrValLysAspGlnGlyLeuSerSerThrSerThrLe	707
Db	1980	GACTTATGTATTACCTTTGACTGTCAAAGATGAGAGGAACCTACAGACGACAGCTCCGT	2039
Qy	707	uThrValAlaValLysLysGluAenAenSerProProArgAlaAraGlaGlyArgHi	727
Db	2040	TAATGTCTATTGTCAAGAGAATAAACAAACCGCCAGTAGCCACAGATCGCTGGGAACGT	2099
Qy	727	sValLeuValLeuProAenAenSerIleThrLeuAspGlySerArgSerThrAspAspGl	747
Db	2100	GGTGGTGACCTTGCCACGACGACAGCTGAGCTGGATGGCTCGAGGTCTCTCAGATGACAA	2159
Qy	747	nArgIleValSerTyrLeuTrpIleArgAspGlyGlnSerProAlaAlaGlyAspValIl	767
Db	2160	GGGGATAGTCAGCTTACCTGTGGACTCGAGATGAGACGAGCCGACGCGAGGGAGTGTCT	2219
Qy	767	eAspGlySerAspHisSerValAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrTh	787
Db	2220	GAATCACTCTGACCAACACCCCTCTCTCTCTCTCTCCACCTGTGTGAGGGGACCTTACAC	2279
Qy	787	rPheHisLeuArgValThrAspSerGlnGlyAlaSerAspThrAspThrAlaThrValGl	807
Db	2280	GTITTCACCTGAAAGTGACAGATCAAAGGGCGAGAGCGACACAGACGCGACGACAGTGGGA	2339
Qy	807	uValGlnProAspProArgLysSerGlyLeuValGluLeuThrLeuGlnValGlyValGl	827
Db	2340	AGTGAACCCCTGACCCCAAGAAAACAACTTGTGGAGATCATCTTCGATGTGAACCTCAG	2399
Qy	827	yGlnLeuThrGluGlnArgLysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnVa	847
Db	2400	TCAGCTGACTGTAGAGGCTGAAGGGATGTCTATCCGCCAGATTTGGGGTCTCTCTGGGGGT	2459
Qy	847	lLeuAspSerAspIleLysValGlnLysIleArgAlaHisSerAspLeuSerThrValIl	867
Db	2460	GCTGGATTCCGACATCATTTGTCAAAAGATTTCAGCCGTACACGAGCAGACGACCAAGAT	2519
Qy	867	eValPheTyrValGlnSerArgProPheLysValLeuLysAlaAlaGluValAlaAr	887
Db	2520	GTGTTTTTTTGTTCAGAAACAGCCCTCCCAACAGCTCTTTCAAAGGCCATGAGTGGCAGC	2579

QY 887 gAsnLeuHisMetArgLeuSerLysGluLysAlaAspPheLeuLeuPheLysValLeuAr 907
 Db 2580 CATGCTCAAGACGAGCTGCAGAACAGAGGCTACCTCTCTCATCTTCAGAGCCCTGGA 2639
 QY 907 gValAspThrAlaGlyCysLeuLeuLysCysSerGlyHisGlyHisCysAspProLeuTh 927
 Db 2640 AATCAGCACAGTCACATGTCAGCTGAACCTGTTCTGACCATGCCACTGTGACTCATTCAC 2699
 QY 927 rLysArgCysLleCysSerHisLeuTrpMetGlnAsnLeuLleGlnArgTyrIleTrpAs 947
 Db 2700 CAAGCGCTGTGTCTGAGCCCTTTGGATGGAGAAATTCATCAAGGTGCAGCTGAGGGA 2759
 QY 947 pGlyGluSerAsnCysGluTrpSerIlePheTyrValThrValLeuAlaPheThrLeuIl 967
 Db 2760 TGGACACAGCAACTGTGAATGAGAGGTGCTTACGTCTCATCTGCTCTCTTGTCTATGT 2819
 QY 967 eValLeuThrGlyGlyPheThrTrpLeuCysIleCysCysLysAspArgGlnLysArgTh 987
 Db 2820 TGTTCCTTGGGGATCCTGTCATGACTACAACTGCTGCTGCAAGAGGCAAAAA--GG 2876
 QY 987 rLysLleArgLysLysThrLysTyrThrIleLeuAspAsnMetAspGluGlnLysArgMe 1007
 Db 2877 AAAACCCCAAGAGGAAAGCAGATACAAAGATCCTGGATGCCAGAT--CAGGAGAGCCT 2933
 QY 1007 tGluLeuArgPro-----LysTyrGlyIleLysHisArgSerThrGluHisAsnSerSe 1025
 Db 2934 GGAGCTGAACCAACCTCCGAGGAGGAGCAACAGAAAGCCCAAGCTGACGAGCAG 2993
 QY 1025 rLeuMetValSerGluSerGluPheAspSerAspGlnAspThrIlePheSerArgGluLy 1045
 Db 2994 CCGTATGCACTTCTGAATCGAGCTGGACAGCGAC--GATGCCATCTTCACATGCCGAGA 3050
 QY 1045 sMetGluArgGlyAsnProLysValSerMetAsnGlySerIleArgAsnGly 1062
 Db 3051 CCGGAGAGAGGGCAAACTACTGTATGTCAGAATGCTCTGTGCCAAAGCGG 3102
 RESULT 9
 ADI21796
 ID ADI21796 standard; cDNA; 3846 BP.
 AC
 XX ADI21796;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Novel human protein cDNA #55.
 XX
 KW forensic; nutritional source; damaged tissue; diseased tissue;
 KW myeloid cell disorder; lymphoid cell disorder;
 KW bone cartilage tissue growth; tendon tissue growth;
 KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
 KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human;
 KW ss; gene.
 XX
 OS Homo sapiens.
 XX
 PN WO2003025148-A2.
 XX
 PD 27-MAR-2003.
 XX
 XX 19-SEP-2002; 2002WO-US029964.
 XX
 XX 19-SEP-2001; 2001US-0323739P.
 PR 13-SEP-2002; 2002US-00323739.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
 PI Haley-Vicente D;
 XX
 XX WPI; 2003-354603/33.
 DR P-PSDB; ADI21080.

XX New polynucleotides and secreted proteins, useful for treating myeloid or
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
 PT tissue growth or regeneration, in wound healing, and in tissue repair and
 PT replacement.
 XX Claim 1; SEQ ID NO 55; 156pp; English.
 PS
 XX The invention relates to an isolated polynucleotide encoding a
 CC polypeptide with biological activity. The polynucleotide encodes and
 CC polypeptides are useful in diagnostics, forensics, gene mapping,
 CC identification of mutations responsible for genetic disorders and other
 CC traits, to assess biodiversity, as nutritional sources or supplements.
 CC The polynucleotides may also be used as molecular weight markers,
 CC chromosome markers or map related gene positions, or as an antigen to
 CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
 CC useful for raising antibodies, as markers for tissues in which the
 CC corresponding polypeptide is expressed, for re-engineering damaged or
 CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
 CC bone cartilage, tendon, ligament and/or nerve tissue growth or
 CC regeneration, in wound healing, in tissue repair and replacement, in
 CC healing of burns, incisions and ulcers, and in treating cancer. The
 CC present sequence represents cDNA encoding a novel human protein.
 XX
 SQ Sequence 3846 BP; 1008 A; 1005 C; 991 G; 842 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.02e-136 Length: 3846
 Score: 2404.00 Matches: 522
 Percent Similarity: 60.13% Conservative: 134
 Best Local Similarity: 47.85% Mismatches: 318
 Query Match: 43.08% Indels: 117
 DB: 10 Gaps: 17
 US-10-764-390-3 (1-1072) x ADI21796 (1-3846)
 QY 4 ProThrGlyValLeuSerSerLeuLeuLeu-----LeuValThrIleAlaGlyCys-Al 21
 Db 280 CCTGTACTCTGTTTATCTACTGCTTTTGTCTCAGCGTCTCTGTTGTCAACAGATGCCAG 339
 QY 21 aArgLysGlnCysSerGluGlyArgThrTyrSerAsnAlaValIleSerProAsnLeuGl 41
 Db 340 TCAGAGCAGGTGCCAGCAGGCGGGAAGACACAAATTTGGAGTTGGCCTCGAGATCTGGGGGAGA 399
 QY 41 uThrThrArgIleMetArgValSerHisThrPheProValValAlaAspCysThrAlaAlaCy 61
 Db 400 A---AATCACCTCTGGCTTCTTGAAGGAAACCCCTCTCTCAGTCAGATGTTGGCTCGCTG 456
 QY 61 sCysAspLeuSerSerCysAspLeuAlaTrpPheGluGlyArgCysTyrLeuValse 81
 Db 457 CTGCCAGGACTCTGCCTGCCATGCTTTTGGTGGCTAGAGGATGTGCATTTCAGGCAGA 516
 QY 81 rCysProHisLysGluAsnCysGluProLysLysMetGlyProLleArgSerTyrLeuTh 101
 Db 517 CTGCAGCAGGCCCCAGAGCTGCCGGGCTTTTAGGACACACTCTCTCCAATTCATCTGCTGT 576
 QY 101 rPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetMetle 121
 Db 577 GTTT---TTAAAAAAATTCCAACTGCGAGATGATTG----- 610
 QY 121 uAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAspIleArgLysAsple 141
 Db 611 -----GGCTTTTCTACCTGAAGATGATGTACCATCTCT 642
 QY 141 u***PheLeuGlyLysAspTrpGlyLeuGluGluMetSerGluTyr***AspAspTyrAr 161
 Db 643 TCTGGGGCTAGGTGGAACTGGGCA----- 667
 QY 161 gGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAlaGl 181
 Db 668 -----TCTGGAGGCGAGAGCCCAAGAGCTGCACT 699
 QY 181 uTyrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGlyAs 201

Db 700 C----- 700
Qy 201 pSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAsnGluSe 221
Db 701 -AGACCTGCTGTATCTCCAGTACCGACGAG----- 730
Qy 221 rAlaSerThrProAlaProLysLeuProGluArg-----SerValLeuLeuPr 237
Db 731 -----AGCTTAATCAGAAAGTCTCAGAAAGAGAGGTAGTCCCGAGTACGAGTAGTATACACC 783
Qy 237 oLeuProThrThrProSerSerGlyGluValLeuGluLysAlaSerGlnLeuGl 257
Db 784 TATAGTGACAGCATCT-----AAAGTGAATGACTCCCAAGCAATTAGTGTGCTGAC 837
Qy 257 nGluGlnSerSerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProPr 277
Db 838 TACCAGTGGCTCTGCAGAGGTCACCAAGCGATTACAAATTTCCAGTCCCTTAACCCACAGA 897
Qy 277 oAlaSerLeuLeuLeuSer-----SerValThrValGluLysSerPr 291
Db 898 CCTGACTGCAGAGCTGTCTGGTGGCCAAAGAATGTATCAGTGCACCAACCTGAAATATCAGA 957
Qy 291 oValLeuThrValThrProGlySer-----ThrGluHisSe 303
Db 958 GGGTCTTGCTACTACGCCACGACCTCAACAAGTAAAGTTCTGAGAAAACCCAGATTGC 1017
Qy 303 rIleProThrPro-----ProThrSerAlaAlaProSe 314
Db 1018 TGTCCCCAGCCAGTGGCTCCTCTACAGTTATGTACTACCCCTACCCCGCCAGGCTCTTT 1077
Qy 314 rGluSerThrProSerGluLeuProIleSerProThrThrAlaProArgThrValLysGl 334
Db 1078 CCAGAGCACCTCAGCACCATACCCAGTT-----ATAAAGGA 1113
Qy 334 uLeuThrValSerAlaGlyAspAsnLeuIleIleThrLeuProAspAsnGluValGluLe 354
Db 1114 ACTGGTGTATCTCTGGAGAGGTCCAGATACCCCTGCCTAAGAAATGAATGATCAATT 1173
Qy 354 uLysAlaPheValAlaProAlaProValGluThrThrTyrAsnTyrGluTrpAsnLe 374
Db 1174 AAATGCATATGTTCTCCAAGAACCCACCTAAAGGAGAAACCTACACCTACGACTGGCAGCT 1233
Qy 374 uIleSerHisProThrAspTyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeuAs 394
Db 1234 GATTACTCATCTAGAGACTACAGTGGAGAAATGGAAGGAAACATTTCCAGATCTCTCAA 1293
Qy 394 nLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAsnAl 414
Db 1294 ACTATCGAAGCTCACTCCAGGCCTGTATGAATTCAAAGTGATTTAGAGGGTCAAAATGC 1353
Qy 414 aPheGlyGluGlyPheValAsnValThrValLysProAlaArgValAsnLeuProPr 434
Db 1354 CCATGGGGAAGGCTATGTGAACGTGACAGTCAAGCCAGAGCCCGGTAAGATCGGCCCCC 1413
Qy 434 oValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIl 454
Db 1414 CATTGCTATTGTGTCACTCAGTTCAGAGATCTCTTTGGCAACCACTTCTACAGTCAT 1473
Qy 454 eAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHisTyrGluGluLeAs 474
Db 1474 TGATGGCAGTCAAAAGCACTGATGATGAATAATCGTTCCAGTACCATTGGGAAGAACTTAA 1533
Qy 474 nGlyProPheIleGluGluLysThrSerValAspSerProValLeuArgLeuSerAsnLe 494
Db 1534 GGGGCTCTAAGAGAAAGAGATTTCTGAAGATACAGCCATATTAACCTAAGTAAACT 1593
Qy 494 uAspProGlyAsnTyrSerPheArgLeuThrValThrAspSerAspGlyAlaThrAsnSe 514
Db 1594 CGTCCCTGGGAATACACTTTTCAGCTTGACTGTAGTACTCTGATGAGCTACCAACTC 1653
Qy 514 rThrThrAlaAlaLeuIleValAsnAsnAlaValAspTyrProProValAlaAsnAlaGl 534

Db 1654 TACTACTGCAAACTGACAGTGAACAAAGCTGTGGATTACCCCTCTGGCCCAACGACGAG 1713
Qy 534 yProAsnHisThrIleThrLeuProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSe 554
Db 1714 CCCCACCAAGTGTATCATCCCTGCCCCAAAACTCCATCACCTCTTTGGGAACCAAGACAC 1773
Qy 554 rAspAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHi 574
Db 1774 TGATGATCATGGCATCACAGCTATAGTGTGTCTACTCAGCCCAAGCAGCAAGGGAAGT 1833
Qy 574 sValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAs 594
Db 1834 GGTGAGATGCAGGGTGTAGAACACCAACCTTACAGCTCTCTCGATGCAAGAAGGAGA 1893
Qy 594 pTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla***ValTh 614
Db 1894 CTACACTTACAGCTCAGAGTACTGACACATAGGACAGCAGGCGCACTGTCTCAAGTGAC 1953
Qy 614 rValIleValGlnProGluAsnAsnArgProProValAlaValAlaGlyProAspLysGl 634
Db 1954 TGTATTGTGCAACCTGAAACCAATAAGCCTCTCAGCGAGATCGAGCCCGAGATAAGA 2013
Qy 634 uLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSerAspAspHisGl 654
Db 2014 GCTGACCTTCTCTGTGATAGCACCAACCTGGATGGCAGCAAGAGCTCAGATGATCAGAA 2073
Qy 654 yIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValGluMetGluAsnIl 674
Db 2074 AATTATCTCATCTCTCGGAAAAAACACAGGGACCTGATGGGGTGCAGCTCGAGAAATGC 2133
Qy 674 eAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuTh 694
Db 2134 TAACAGCAGTGTCTACTGTGCTGGCTGCAAGTGGGGACCTATGTGTTCACTCTGAC 2193
Qy 694 rValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAlaValLysLysGl 714
Db 2194 TGTCAAGATGAGAGGAACCTGCAAGCCAGAGCTCTGTGAATCTCATTTGTCAGGAAGA 2253
Qy 714 uAsnAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsnAs 734
Db 2254 AATAAAACAACCACTATAGCAAGATAACTGGGAATGTGTGTGATTACCTTACCACGAG 2313
Qy 734 nSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrLeuTr 754
Db 2314 CACAGCAGCTGGATGGCTCTAAGTCTCTCAGATGACAAGGGAATAGTCACTACCTCTG 2373
Qy 754 pIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerVa 774
Db 2374 GACTCGAGATGAGGGGAGCCAGCAGCAGGGAGGTGTAAATCACTCTGACCATCACCC 2433
Qy 774 lAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThrAs 794
Db 2434 TATCTTTTCTTTCAAACTGGTTGAGGGAACCTACATTTTCACTGAAAGTGACCGA 2493
Qy 794 pSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArgLys 814
Db 2494 TGCAAGGGTGAGAGTGACACAGCCGACCACTGTGGAGGTGAACCTGTATCCAGGAA 2553
Qy 814 sSerGlyValGluLeuLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLys 834
Db 2554 AAACAACCTGGTGAGATCATCTTTGGATATCAACGTCAGTCAGCTAAGTGAAGGCTGNA 2613
Qy 834 sAspThrLeuValArgGlnLeuAlaValLeuAsnValLeuAspSerAspIleLysVa 854
Db 2614 GGGGATGTTCACTCCCGCAGATTGGGGTCTCTCTGGGGTGTGTGATTCCGACATCATGT 2673
Qy 854 lGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTyrValGlnSerAr 874
Db 2674 GCAAAAGATTTCAGCCGTACACGGGACAGCACCAAAATGGTATTTTTTGTGTAACACGA 2733
Qy 874 gProProPheLysValLeuLysAlaAlaGluValAlaArgAsnLeuHisMetArgLysSe 894
Db 2734 GCCTCCCCACCATCTTTCAAGGGCCATGAGGTGGCAGCGATGTCTCAAGATGAGCTGCG 2793

533	Db		CTGCAGCAGCCCGCCAGAGCTGCCGGGCTTTTAGGACACACACTCTCTCCAAATTCATGCTGGT	592
101	Qy		rPheValLeuArgProValGlnArgProAlaGlnLeuLeuAspTyrGlyAspMetLe	121
593	Db		GTIT--TTAAANAATTCACAACTGCAGATGTTG-----	626
121	Qy		uAsnArgGlySerProSerGlyIleTrpGlyAspSerProGluAspIleArgLysAspLe	141
627	Db		-----GGCTTTCTACTCTGAAGATGATGTACCACATCT	658
141	Qy		uPheLeuGlyLysAspTrpGlyLeuGluMetSerGluTyr***AspAspTyrAr	161
659	Db		TCTGGGCTAGGTGGAACTGGCA-----	683
161	Qy		gGluLeuGluLysAspLeuLeuGlnProSerGlyLysGlnGluProArgGlySerAlaGl	181
684	Db		-----TCTTGAGGCAGAGCCGCCAGAGCTGCACCT	715
181	Qy		uTyrThrAspTrpGlyLeuLeuProGlySerGluGlyAlaPheAsnSerSerValGlyAs	201
716	Db		C-----	716
201	Qy		pSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAsnGluSe	221
717	Db		-AGACCTGCTGTATCTTCCAGTGACCCAGCAG-----	746
221	Qy		rAlaSerThrProAlaProLysLeuProGluArg-----SerValLeuLeuPr	237
747	Db		-----AGCTTAATCAGGAAGCTTCAGAGAGAGGTAGTCCCACTGACGTAGTTACAC	799
237	Qy		oLeuProThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGl	257
800	Db		TATAGTCACACAGCATCT-----AAAGTGAATGACTCCAACTAGTGGTCTGCAC	853
257	Qy		nGluGlnSerSerAsnSerSerGlyLysGluValLeuMetProSerHisSerLeuProPr	277
854	Db		TACCAGTGGCTCTGCAGAGGTCCCAAGGCGATTACAAATTTCCAGTCCCTTAACCCACAGA	913
277	Qy		oAlaSerLeuLeuLeuSer-----SerValThrValGluLysSerPr	291
914	Db		CCTGACTGCAGAGCTGTCTGGTGGCCAAAGAATGTATCAGTGCACCTGAATATCAGA	973
291	Qy		oValLeuThrValThrProGlySer-----ThrGluHisSe	303
974	Db		GGGTCTTGCTACTACGCCAGCACTCAACAAGTAAAGTTCTGAGAAAACCCAGATGTC	1033
303	Qy		rIleProThrPro-----ProThrSerAlaAlaProSe	314
1034	Db		TGTCCTCCCGCAGCAGTGGCTCCTCTACAGTTATGCTACCCCTACCCCTCCAGGCTCTTT	1093
314	Qy		rGluSerThrProSerGluLeuProLysSerProThrThrAlaProArgThrValLysGl	334
1094	Db		CCAGGCACCTTCAGCACAATACCCAGTT-----ATPAAAGGA	1129
334	Qy		uLeuThrValSerAlaGlyAspAsnLeuIleThrLeuProAspAsnGluValGluLe	354
1130	Db		ACTGGTGGTATCTCTGGAGAGAGTGTCAGATAACCTGCCTGAAGATGAAGTTCAATT	1189
354	Qy		uLysAlaPheValAlaProAlaProProValGluThrThrTyrAsnTyrGluTrpAsnLe	374
1190	Db		AAATGTCATATGTTCTCCAAGAACCACTTAAGAGGAGAAACCTTACACCTACGACTGGCAGCT	1249
374	Qy		uIleSerHisProThrAspTyrGlnGlyGluLysGlnGlyHisLysGlnThrLeuAs	394
1250	Db		GATTACTCATCTTAGAGACTACAGTGAGAGAAATGGAAGGGGAAACATTTCCCGATCCTCAA	1309
394	Qy		nLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAsnAl	414
1310	Db		ACTATCAAGAGCTCACTCCAGGCTGTATGAATTCAAAGTGATTGTAGAGGTCAAAATGC	1369
414	Qy		aPheGlyGluGlyPheValAsnValThrValLysProAlaArgValAsnLeuLeuProPr	434

Db 717 C----- 717
Qy 201 pSerProAlaValProAlaGluThrGlnGlnAspProGluLeuHisTyrLeuAenGluSe 221
Db 718 -AGACCTGCTGTATCTCCAGTGACCGAG- 747
Qy 221 rAlaSerThrProAlaProLysLeuProGluArg-----SerValLeuLeuPr 237
Db 748 -----AGCTTAATCAGGAAGCTTCAGAAAGAGAGTAGTCCCGAGTGACGTAGTATACAC 800
Qy 237 oLeuProThrThrProSerSerGlyGluValLeuGluLysGluLysAlaSerGlnLeuGl 257
Db 801 TATAGTGACACAGTATCT-----AAAGTGAATGACTCCCAACGAATTAGTGCTGAC 854
Qy 257 nGluGlnSerSerAenSerSerGlyLysGluValLeuMetProSerHisSerLeuProPr 277
Db 855 TACCAGTGGCTCTCGCAGAGGTCACACAGGCGAATACAAATTTCCAGTCCCTTAACCCACAGA 914
Qy 277 oAlaSerLeuLeuLeuSer-----SerValThrValGluLysSerPr 291
Db 915 CTTGACTGCAGAGCTGTCTGGTGGGCCCAAGAAATGTATCAGTGCACCACTGAAATATCAGA 974
Qy 291 oValLeuThrValThrProGlySer-----ThrGluHisSe 303
Db 975 GGGTCTTGCTACTAGCCACGACTCAACAAGTAAAGTTCTGAGAAAACCCAGATTGC 1034
Qy 303 rIleProThrPro-----ProThrSerAlaAlaProSe 314
Db 1035 TGTCCCCCAGCAGTGGCTCCTCTACAGTATTATGCTACCCCTACCCCCAGGCGCTCTTT 1094
Qy 314 rGluSerThrProSerGluLeuProLysSerProThrThrAlaProArgThrValLysGl 334
Db 1095 CCAGACACCTTCAGCACCATCCAGTT-----ATAAGGA 1130
Qy 334 uLeuThrValSerAlaGlyAspAenLeuIleIleThrLeuProAspAenGluValGluLe 354
Db 1131 ACTGGTGTATCTGCTGGAGAGAGTGCTCAGATTAACCTCGCTAAGAAATGAATTCAAAT 1190
Qy 354 uLysAlaPheValAlaProAlaProValGluThrThrTyrAenTyrGluTrpAenLe 374
Db 1191 AAATGCATATGTTCTCCAAGAACCCACCTAAAGGAGAAACCTACACCTACGACTGGCAGCT 1250
Qy 374 uIleSerHisProThrAspTyrGlnGlyGluIleLysGlnGlyHisLysGlnThrLeuAs 394
Db 1251 GATTACTCATCTAGACATACAGTGGAGAAATGGAAGGAAACATTTCCAGATCTCCTCAA 1310
Qy 394 nLeuSerGlnLeuSerValGlyLeuTyrValPheLysValThrValSerSerGluAenAl 414
Db 1311 ACTATCGAAGCTCACTCCAGGCCCTGTATGAATTCAAAGTGATGTAGAGGTCAAAATGC 1370
Qy 414 aPheGlyGluGlyPheValAenValThrValLysProAlaArgValAenLeuProPr 434
Db 1371 CCATGGGGAGGCTATGTGAACGTGACAGTCAAGCCAGAGCCCGCTAAGATCGGCCCCC 1430
Qy 434 oValAlaValValSerProGlnLeuGlnGluLeuThrLeuProLeuThrSerAlaLeuIl 454
Db 1431 CATTGCTATTGTGCACCTCAGTTCAGGAGATCTTTGGCAACCACTTCTACAGTCAT 1490
Qy 454 eAspGlySerGlnSerThrAspAspThrGluIleValSerTyrHisTTPGluGluIleAs 474
Db 1491 TGATGGCAGTCAAAAGCACTGAATGATGAATAATCGTTACGATTCATGGGAAGAACTTAA 1550
Qy 474 nGlyProPheIleGluGluLysThrSerValAspSerProValLeuArgLeuSerAenLe 494
Db 1551 GGGGCCCTCTAAGAGAAGAGAGATTTCTGAAGATACAGCCATATTAACCTAAGTAACCT 1610
Qy 494 uAspProGlyAenTyrSerPheArgLeuThrValThrAspSerAspGlyAlaThrAenSe 514
Db 1611 CGTCCCTGGGAACCTACACTTTTCAGCTTGACTGTAGTACTCTGATGAGCTACCAACTC 1670
Qy 514 rThrThrAlaAlaLeuIleValAenAenAlaValAspTyrProProValAlaAenAlaGl 534

Db 1671 TACTACTGCAAACTGACAGTGAACAAAGCTGTGGATTACCCCTCTGGCCAAACGACGG 1730
Qy 534 yProAenHisThrIleThrLeuProGlnAenSerIleThrLeuAenGlyAenGlnSerSe 554
Db 1731 CCCCAACCAAGTGTATCACCTGCCCCCAAACTCCACCAACCTCTTTGGGAACCAAGACAC 1790
Qy 554 rAspAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLysHi 574
Db 1791 TGATGATCATGGCATCACAGCTATGATGTGTCTACTAGCCCAAGCAGCAAGGAAAGT 1850
Qy 574 sValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAs 594
Db 1851 GGTGAGATGCAGGGTGTAGAACCAACCACTTACAGCTCTCTCGATGCAAGAAGGAGA 1910
Qy 594 pTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla***ValTh 614
Db 1911 CTACACTTCACAGCTCAGAGTGACTGACACAAATAGGACAGCAGGCGCACTGCTCAAGTGC 1970
Qy 614 rValIleValGlnProGluAenAenArgProProValAlaValAlaGlyProAspLysGl 634
Db 1971 CGTTATTGTGCAACCTGAAACCAATAAGCCCTCTCAGCGAGATCGAGCCCGAGATAAAGA 2030
Qy 634 uLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSerAspAspHisGl 654
Db 2031 GCTGACCTTCTGTGTAGTACACCACTCGATGGCAGCAAGAGCTCAGATGATCAGAA 2090
Qy 654 yIleValPheTyrHisTrpGlu-HisValArgGlyProSerAlaValGluMetGluAenI 674
Db 2091 AATTATCTCATATCTCTGGGAAAAAAACACAGGACCTGATGGGTGTCAGCTCGAATG 2150
Qy 674 leAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeuT 694
Db 2151 CTAAACAGCAGTGTGTACTGTGTGCTGGCTGCAAGTGGGACCTATGTGTTCACCTTGA 2210
Qy 694 hrValLysAspGlnGlnGlyLeuSerSerThrSerThrLeuThrValAlaValLysLysG 714
Db 2211 CTGTCAAGATGAGAGAACTTGCAAAGCCAGAGCTCTGTGAATGTCTATTGTCAAGAAG 2270
Qy 714 luAenAenSerProProAlaArgAlaGlyArgHisValLeuValLeuProAenA 734
Db 2271 AAATAAAACAAACCACTATAGCCAAGATACTGGGAATGTGGTATACCTACCCACGA 2330
Qy 734 snSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrLeuT 754
Db 2331 GCACAGCAGAGCTGGATGGCTCTAAGTCTCAGATGACAAAGGAATAGTACGTACCTCT 2390
Qy 754 rPileArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSerV 774
Db 2391 GGACTCGAGATGAGGGAGGCCAGCAGCAGGGGAGGTGTAAATCACTCTGACCATCACC 2450
Qy 774 alAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThrA 794
Db 2451 CTATCCTTTTCTTTTCAAACTGTGTTGAGGAACTTACACTTTTTCACCTGAAAGTGACCG 2510
Qy 794 spSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArgL 814
Db 2511 ATGCAAGGGTGAAGTGAACAGACAGCCAGCCACTGTGGAGGTGAAACCTGATCCCAAG 2570
Qy 814 ysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgL 834
Db 2571 AAAACAACTGGTGGAGATCATCTTTGGATATCAACGTCAGTCAGCTAACTGAGAGGCTGA 2630
Qy 834 ysAspThrLeuValArgGlnLeuAlaValLeuLeuAenValLeuAspSerAspIleLysV 854
Db 2631 AGGGGATGTTTCATCCGCGCAGATTTGGGTCTCTCTGGGGGTGCTGATTTCCGACATCATG 2690
Qy 854 alGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTyrValGlnSerA 874
Db 2691 TGCAAAAAGATTCAAGCCGTACACGAGCAGACCAAAAATGGTATTTTGTTCAAAACG 2750
Qy 874 rgProProPheLysValLeuLysAlaAlaGluValAlaArgAenLeuHisMetArgLeuS 894
Db 2751 AGCCTCCCAACGATCTTCAAGGCCCATGAGGTGGCAGCGGATGCTCAAGAGTGAGCTGC 2810

Qy 190 GlySerGluGlyAlaPheAsnSerSerValGlyAspSerProAlaValProAlaGluThr 209
Db 732 -----AGACCTGCTGTATCTTCCAGTGAC 755
Qy 210 GlnGlnAspProGluLeuHisTyrLeuAsnGluSerAlaSerThrProAlaProLysLeu 229
Db 756 CAGCAG-----AGCTTAATCAGGAGGTT 779
Qy 230 ProGluArg-----SerValLeuLeuProLeuProThrThrProSerSerGly 245
Db 780 CAGAGAGAGGAGGTAGTCCAGTCAGCTAGTATACACCTATAGTGACAGCAATCT----- 833
Qy 246 GlnValLeuGluLysGluLysAlaSerGlnLeuGlnGlnSerSerAsnSerSerGly 265
Db 834 AAGTGAATGACTCCAAAGAAATAGGTGGTCTGACTACAGTGGCTCTGCAGAGGTCCAC 893
Qy 266 LysGluValLeuMetProSerHisSerLeuProProAlaSerLeuGluLeuSer----- 283
Db 894 AAGCGATTACAATTCAGTCCCTTAACACAGACCTGACTGCAGAGCTGTCTGGTGG 953
Qy 284 -----SerValThrValGluLysSerProValLeuThrValThrProGlySer 299
Db 954 CCAAAGAATGTATCAGTCAACCTGAAATATCAGAGGGTCTTGCTACTACGCCCGACACT 1013
Qy 300 -----ThrGluHisSerIleProThrPro----- 307
Db 1014 CAACAAGTAAAGTTCTGAGAAACCCAGAAATGCTGTCCTCCAGCCAGTGGCTCCCTCC 1073
Qy 308 -----ProThrSerAlaAlaProSerGluSerThrThrProSerGluLeuPro 322
Db 1074 TACAGTTATGCTACCCCTACCCCGAGGCTCTTTCAGAGACCTCAGCACCATACCCA 1133
Qy 323 IleSerProThrThrAlaProArgThrValLysGluLeuThrValSerAlaGlyAspAsn 342
Db 1134 GTT-----ATAAGGAACCTGGTGTATCTGCTGGAGAGAGT 1169
Qy 343 LeuIleIleThrLeuProAspAsnGluValGluLeuLysAlaPhe-ValAlaProAla-P 362
Db 1170 GTCCAGATTAACCTGCTTGAAGATGAAGTTCAATTAATGATATGTTCTTCCAGGAAC 1229
Qy 362 roProValGluThrThrTyrAsnTyrGluTTPAsnLeuIleSerHisProThrAspTyrG 382
Db 1230 CACCTAAAGGAGAAACCTACACCTACGACTGCGAGCTGATTCATCTAGAGACTACA 1289
Qy 382 InGlyGluIleLysGlnGlyHisLysGlnThrLeuAsnLeuSerGlnLeuSerValGlyL 402
Db 1290 GTGGAGAAATGGAAGGGAACATTCCAGATCTCAAACTATCGAAGCTCACTCCAGGCC 1349
Qy 402 euTyrValPheLysValThrValSerSerGluAsnAlaPheGlyGluGlyPheValAsnV 422
Db 1350 TGTATGAATTCAAAGTGATTTAGAGGGTCAAAATGCCCATCGGGAAGGCTATGTGAACG 1409
Qy 422 alThrValLysProAlaArgValAsnLeuProProValAlaValValSerProGlnL 442
Db 1410 TGACAGTCAAGCCAGGCCCCGTAGAATCGCCCCCATTTGCTATTGTGTCACCTCAGT 1469
Qy 442 euGlnGluLeuThrLeuProLeuThrSerAlaLeuIleAspGlySerGlnSerThrAspA 462
Db 1470 TCCAGGAGATCTCTTTGCCAACCACTTCTACAGTCAATTCATGGCAGTCAAAAGCACTGATG 1529
Qy 462 spThrGluIle-ValSer-TyrHisTTPGluGluIleAsnGlyProPheIleGluLys 481
Db 1530 ATGATAAAATCGTTTCAGTACCATTTGGGAAGAACTTAAGGGGCGCTCTTAAGAGAAGAGAG 1589
Qy 482 ThrSerValAspSerProValLeuArgLeuSerAsnLeuAspProGlyAsnTyrSerPhe 501
Db 1590 ATTTCTGAAGTACAGCCATATTAATAAGTAAACTCGCCCCCTGGGAACTACACTTTC 1649
Qy 502 ArgLeuThrValThrAspSerAspGlyAlaThrAsnSerThrThrAlaAlaLeuIleVal 521
Db 1650 AGCTTGACTAGTAGACTCTGTAGAGCTACCAACTCTACTACTGCAAACTGACAGTG 1709

Qy 522 AsnAsnAlaValAspTyrProProValAlaAsnAlaGlyProAsnHisThrIleThrLeu 541
Db 1710 AACAAAGCTGTGGATTAACCCCTGTGGCCCAACGAGCGCCCAACCAAGTATCACCCCTG 1769
Qy 542 ProGlnAsnSerIleThrLeuAsnGlyAsnGlnSerSerAspAspHisGlnIleValLeu 561
Db 1770 CCCCAAAATCCATCACCTCTTTTGGGAACCAGAGCACTGATGATCATGGCATCACCAAGC 1829
Qy 562 TyrGluTTPSerLeuGlyProGlySerGluGlyLysHisValValMetGlnGlyValGln 581
Db 1830 TATCAGTGGTCACTCAGCCCAAGCAGCAAGGGAAAGTGGTGGAGATGCAGGGGTGTAGA 1889
Qy 582 ThrProTyrLeuHisLeuSerAlaMetGlnGluGlyAspTyrThrPheGlnLeuLysVal 601
Db 1890 ACACCAACCTTACAGCTCTCTCGATGCAAGAAGAGACTACACTTACCACCTCACAGTG 1949
Qy 602 ThrAspSerSerArgGlnGlnSerThrAla***ValThrValIleValGlnProGluAsn 621
Db 1950 ACTGACACAATAGCAGACAGCGCCACTGCTCAAGTGACTGTATTATTGTGCAACCTGAAAC 2009
Qy 622 AsnArgProProValAlaValAlaGlyProAspLysGluLeuIlePheProValGluSer 641
Db 2010 AATAAGCTCTCTCAGGAGATGCAGGCCCAAGATAAAGAGCTGACCTTCTCTGTGGATAGC 2069
Qy 642 AlaThrLeuAspGlySerSerSerSerAspAspHisGlyIleValPheTyrHisTTPGlu 661
Db 2070 ACAACCTGGATGCGAGAGAGCTCAGATGATCAGAAATTTATCTCATATCTCTGGGA 2129
Qy 662 HisValArgGlyProSerAlaVal-GluMetGluAsnIleAspLysAlaIleAlaThrVa 681
Db 2130 AAAACACAGGAGCTGATGGGTCTATCTCGAGAATGCTAACACAGCGTGTCTACTGT 2189
Qy 681 lThrGlyLeuGlnValGlyThrTyrHisPheArgLeuThrValLysAspGlnGlnGlyLe 701
Db 2190 GACTGGGTGCGCAGTGGGACCTATGTGTTCACCTTGACTGTCAAGATGATAGGAACCT 2249
Qy 701 uSerSerThrSerThrLeuThrValAlaValLysLysGluAsnAsnSerProProArgAl 721
Db 2250 GCAAAGCTGAGCTCTCTGTAATGTCATTGTCCAAGAACAAATAAACCAACACCTATAGC 2309
Qy 721 aArgAlaGlyLysArgHisValLeuValLeuProAsnAsnSerIleThrLeuAspGlySe 741
Db 2310 CAAGATACTGGGAATGTGGTGATTACCTTACCACCGACACATCAGAGCTGGATGGCTC 2369
Qy 741 rArgSerThrAspAspGlnArgIleValSerTyrLeuTTPIleArg-AspGlyGlnSerP 761
Db 2370 TAAGACCTCAGATGACCAAGGATAGTCAGTACTCTGGACTCGAAAGAGGGGAGGCC 2429
Qy 761 roAlaAlaGlyAspValIleAspGlySerAspHisSerValAlaLeuGlnLeuThrAsnL 781
Db 2430 CAGCAGAGGGGAGGTGTAAATCACTTGACCATCACCTTATCTCTTTTCTTTCAAAACC 2489
Qy 781 euValGluGlyValTyrThrPheHisLeuArgValThrAspSerGlnGlyAlaSerAspT 801
Db 2490 TGGTTGAGGNAACCTACATCTTTCACCTGAAAGTGACCGATGCAAAAGGTGAGAGTGACA 2549
Qy 801 hrAspThrAlaThrValGluValGlnProAspProArgLysSerGlyLeuValGluLeuT 821
Db 2550 CAGACCGACCACTGTGGAGGTGAACCTGATCCAGGAAAAACAACCTGTGGAGATCA 2609
Qy 821 hrLeuGlnValGlyValGlyGlnLeuThrGluGlnArgLysAspThrLeuValArgGlnL 841
Db 2610 TCTTGGATATCAACGTCAGTCAGCTAGAGAGGTGAAGGGGATGTTTCCTCGCCAGA 2669
Qy 841 euAlaValLeuLeuAsnValLeuAspSerAspIleLysValGlnLysIleArgAlaHis 861
Db 2670 TTGGGGTCTCTCTGGGGTGTGGATTCGACATCATTTGTGCAAAAGATTTCAGCCGTACA 2729
Qy 861 erAspLeuSerThrValIleValPheTyrValGlnSerArgProProPheLysValLeuL 881
Db 2730 CGGAGCAGAGCACCACCAAAATGGTATTTTGTTCAAAACAGAGCCTCCCCACCATCTTTA 2789
Qy 881 ySAlaAlaGluValAlaArgAsnLeuHisMetArgLeuSerLysGluLysAlaAspPheL 901

Db 543 GGGCCCAACCAAGTGATCACCTGCCCCCAAACTCCATCACCTCTTTGGGAACAGAGC 602
 Qy 554 SerAspHisGlnIleValLeuTyrGluTrpSerLeuGlyProGlySerGluGlyLys 573
 Db 603 ACTGATGATCATGGCATCACAGCTATGAGTGGTCACTCAGCCCAAGCAGCAAAAGGGA 662
 Qy 574 HisValValMetGlnGlyValGlnThrProTyrLeuHisLeuSerAlaMetGlnGly 593
 Db 663 GTGTGGNGATGCAAGGGTGTAGAACCAACCTTACAGCTCTCTGCGATGCAAGAAGGA 722
 Qy 594 AspTyrThrPheGlnLeuLysValThrAspSerSerArgGlnGlnSerThrAla**Val 613
 Db 723 GACTACACTTACCAGCTCACAGTGACTGACACAAATAGGACAGAGCCACCTGCTCAAGTG 782
 Qy 614 ThrValIleValGlnProGluAsnAsnArgProProValAlaValAlaGlyProAspLys 633
 Db 783 ACTGTATTGTGCAACCTTCAAAACAATAAAGCTCTCTCAGGCAGATGCAAGGCCAGATAA 842
 Qy 634 GluLeuIlePheProValGluSerAlaThrLeuAspGlySerSerSerSerAspHis 653
 Db 843 GAGCTGACCTTCTGTGGATAGCACAAACCTGTGATGGCAGCAAGCTCAGATGATCAG 902
 Qy 654 GlyIleValPheTyrHisTrpGluHisValArgGlyProSerAlaValGluMetGluAsn 673
 Db 903 AAAATTATCATATCTCTGGGAAAAACACAGAGGACCTGTAGGGTGCACTCGAGAT 962
 Qy 674 IleAspLysAlaIleAlaThrValThrGlyLeuGlnValGlyThrTyrHisPheArgLeu 693
 Db 963 GCTAACAGCAGCTGTGTACTGTGACTGGGTGCAAGTGGGACCTATGTCTTCACTTG 1022
 Qy 694 ThrValLysAspGlnGlnGlyLeuSerSerThrThrLeuThrValAlaValLysLys 713
 Db 1023 ACTGTCAAGATGAGAGGAACCTGCAAGCCAGAGCTCTGTGAATGTCAITGTCAAAGAA 1082
 Qy 714 GluAsnSerProProArgAlaArgAlaGlyGlyArgHisValLeuValLeuProAsn 733
 Db 1083 GAAATAACAACCAACCTATAGCAAGATACTGGGAATGTGGTATACCTTACCCACG 1142
 Qy 734 AsnSerIleThrLeuAspGlySerArgSerThrAspAspGlnArgIleValSerTyrIleu 753
 Db 1143 AGCACAGCAGCTGGATGGCTCTAAGTCTCAGATGACAGGAAGTAAGTACAGTACCTC 1202
 Qy 754 TrpIleArgAspGlyGlnSerProAlaAlaGlyAspValIleAspGlySerAspHisSer 773
 Db 1203 TGGACTCGAGATGAGGGAGGCCAGCAGCAGGGAGGTGTAAATCACTCTGACCATCAC 1262
 Qy 774 ValAlaLeuGlnLeuThrAsnLeuValGluGlyValTyrThrPheHisLeuArgValThr 793
 Db 1263 CCTATCTTTTCTTTCAACCTGGTTGAGGGAACCTACACTTTTTCACCTGAAAGTGACC 1322
 Qy 794 AspSerGlnGlyAlaSerAspThrAspThrAlaThrValGluValGlnProAspProArg 813
 Db 1323 GATCAAAAGGTGAGAGTGACACAGACCGGACCACTGTGGAGGTGAAACCTGATCCACG 1382
 Qy 814 LysSerGlyLeuValGluLeuThrLeuGlnValGlyValGlyGlnLeuThrGluGlnArg 833
 Db 1383 AAAAAACAACCTGTGGAGATCATCTTGGATATCAACGTCAGTCAGTAACCTGAGAGGGTG 1442
 Qy 834 LysAspThrLeuValArgGlnLeuAlaValLeuLeuAsnValLeuAspSerAspIleLys 853
 Db 1443 AAGGGATGTTTCACTCCCGCAGATGGGTCTCTCTGGGGGTGCTGGATCCGATCATTT 1502
 Qy 854 ValGlnLysIleArgAlaHisSerAspLeuSerThrValIleValPheTyrValGlnSer 873
 Db 1503 GTGCAAAAGATTACCGCTACACGAGCAGAGCAGCAACCAAAATGGTATTTTTTTTCAAAC 1562
 Qy 874 ArgProProPheLysValLeuValAlaAlaGluValAlaArgAsnLeuHisMetArgLeu 893
 Db 1563 GAGCCTCCCAACCATGATCTTCAAAAGGCCATGAGGTGGCAGCGATGCTCAAGATGAGGTG 1622
 Qy 894 SerLysGluLysAlaAspPheLeuPheLysValLeuArgValAspThrAlaGlyCys 913
 Db 1623 CGGAAGCAAAAGGCAGACTTTTTTGTATTTACAGACCTTGGGAAGTCAACACTGTACATGT 1682

Qy 914 LeuLeuLysCysSerGlyHisGlyHisCysAspProLeuThrLysArgCysIleCysSer 933
 Db 1683 CAGCTGAACCTGTCGACCACTGTCGACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1742
 Qy 934 HisLeuTrpMetGluAsnLeuIleGlnArgTyrIleTrpAspGlyGluSerAsnCysGlu 953
 Db 1743 CCTTTTGGATGGAGAAATTCATCAAGTGCGAGCTGAGGGATGGAGCAGCAACTGTGAG 1802
 Qy 954 TrpSerIlePheTyrValThrValLeuAlaPheThrLeuIleValLeuThrGlyGlyPhe 973
 Db 1803 TGGAGCGTGTATATGTTATCATTTGCTACCTTTGTCATTTGTTGCTTGGGAATCCTG 1862
 Qy 974 ThrTrpLeuCysIleCysCysCysGlyArgGlnLysArgThrLysIleArgLysLysThr 993
 Db 1863 TCTTGGAGCTGTGATCTGTTGTTGTAAGAGGCAAAA---GGAAAAACCAAGAGGAAAGC 1919
 Qy 994 LysTyrThrIleLeuAspAsnMetAspGluGlnGluArgMetGluLeuArgPro----- 1011
 Db 1920 AAGTACAGATCTCTGGATGCCACGAT---CAGGAAAGCCTGGAGCTGAGCCAACTCC 1976
 Qy 1012 LysTyrGlyIleLysHisArgSerThrGluHisAsnSerSerLeuMetValSerGluSer 1031
 Db 1977 CGAGCAGGCATCAAAACAGAAAGGCTTTTGTCTAAGTAGCAGCTGATGCACTCCGAGTCA 2036
 Qy 1032 GluPheAspSerAspGlnAspThrIlePheSerArgGluLysMetGluArgGlyAsnPro 1051
 Db 2037 GAGCTGGACACGCGAT---GATGCCATCTTTACATGGCCAGACCCGAGAGAGGCAAACTC 2093
 Qy 1052 LysValSerMetAsnGlySerIleArgAsnGly 1062
 Db 2094 CTGCATGTCAGAAATGCTCTGTACCCCAACGGG 2126
 RESULT 15
 ID AAV58761 standard; cDNA; 2501 BP.
 XX AAV58761;
 AC AAV58761;
 DT 18-JAN-1999 (first entry)
 DE Human secreted protein cw1233_3 cDNA.
 KW Secreted protein; human; cw1233_3; db.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 364..780
 FT /*tag= a
 WO9840404-A2.
 PD 17-SEP-1998.
 XX 09-MAR-1998; 98WO-US004601.
 PF 11-MAR-1997; 97US-00815381.
 PR 06-MAR-1998; 98US-00036321.
 PA (GEM) GENETICS INST INC.
 XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
 PI Spaulding V, Agostino MJ;
 DR WPI; 1998-520802/44.
 DR P-PSDB; AAW69431.
 XX New isolated polynucleotides and secreted proteins - useful as, e.g.
 PT nutritional additives, immunostimulators, haemopoiesis regulators and
 PT as diagnostic agents.
 XX Claim 32; Page 88-89; 119pp; English.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:22:50 ; Search time 70 Seconds
(without alignments)
6370.973 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 1072
Sequence: 1 MAPPTGVLSLLLVTTIAGC.....VSMNGSIRNGASFSYCSKDR 1072

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1854112 seqs, 416015017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 458990

Minimum DB seq length: 0

Maximum DB seq length: 35

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	1.7	19	18	US-10-450-763-52425
2	16	1.5	16	16	US-10-764-390-43
3	14	1.3	24	16	US-10-764-390-265
4	14	1.3	29	16	US-10-764-390-262
5	14	1.3	29	16	US-10-764-390-268
6	9	0.8	9	16	US-10-764-390-44
7	9	0.8	9	16	US-10-764-390-45
8	9	0.8	9	16	US-10-764-390-46
9	9	0.8	9	16	US-10-764-390-47
10	9	0.8	9	16	US-10-764-390-48
11	9	0.8	9	16	US-10-764-390-49

12	9	0.8	9	16	US-10-764-390-50	Sequence 50, Appl
13	9	0.8	9	16	US-10-764-390-51	Sequence 51, Appl
14	9	0.8	9	16	US-10-764-390-52	Sequence 52, Appl
15	9	0.8	9	16	US-10-764-390-53	Sequence 53, Appl
16	9	0.8	9	16	US-10-764-390-54	Sequence 54, Appl
17	9	0.8	9	16	US-10-764-390-55	Sequence 55, Appl
18	9	0.8	9	16	US-10-764-390-56	Sequence 56, Appl
19	9	0.8	9	16	US-10-764-390-59	Sequence 59, Appl
20	9	0.8	9	16	US-10-764-390-60	Sequence 60, Appl
21	9	0.8	9	16	US-10-764-390-61	Sequence 61, Appl
22	9	0.8	9	16	US-10-764-390-62	Sequence 62, Appl
23	9	0.8	9	16	US-10-764-390-63	Sequence 63, Appl
24	9	0.8	9	16	US-10-764-390-64	Sequence 64, Appl
25	9	0.8	9	16	US-10-764-390-65	Sequence 65, Appl
26	9	0.8	9	16	US-10-764-390-66	Sequence 66, Appl
27	9	0.8	9	16	US-10-764-390-67	Sequence 67, Appl
28	9	0.8	9	16	US-10-764-390-68	Sequence 68, Appl
29	9	0.8	9	16	US-10-764-390-69	Sequence 69, Appl
30	9	0.8	9	16	US-10-764-390-70	Sequence 70, Appl
31	9	0.8	9	16	US-10-764-390-71	Sequence 71, Appl
32	9	0.8	9	16	US-10-764-390-72	Sequence 72, Appl
33	9	0.8	9	16	US-10-764-390-73	Sequence 73, Appl
34	9	0.8	9	16	US-10-764-390-74	Sequence 74, Appl
35	9	0.8	9	16	US-10-764-390-75	Sequence 75, Appl
36	9	0.8	9	16	US-10-764-390-76	Sequence 76, Appl
37	9	0.8	9	16	US-10-764-390-77	Sequence 77, Appl
38	9	0.8	9	16	US-10-764-390-78	Sequence 78, Appl
39	9	0.8	9	16	US-10-764-390-79	Sequence 79, Appl
40	9	0.8	9	16	US-10-764-390-80	Sequence 80, Appl
41	9	0.8	9	16	US-10-764-390-81	Sequence 81, Appl
42	9	0.8	9	16	US-10-764-390-82	Sequence 82, Appl
43	9	0.8	9	16	US-10-764-390-83	Sequence 83, Appl
44	9	0.8	9	16	US-10-764-390-84	Sequence 84, Appl
45	9	0.8	9	16	US-10-764-390-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-10-450-763-52425
; Sequence 52425, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 52425
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(19)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-52425

Query Match 1.7%; Score 18; DB 18; Length 19;
Best Local Similarity 100.0%; Pred.No. 4.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 MAPPTGVLSLLLVTTIA 18
|||||

```
Db      1  MAPPTGVLSLLLVIA 18

RESULT 2
US-10-764-390-43
; Sequence 43, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-43

Query Match      1.5%; Score 16; DB 16; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      522  NNAVDYPPVANAGPNH 537
Db      1  NNAVDYPPVANAGPNH 16

RESULT 3
US-10-764-390-265
; Sequence 265, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 265
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-265

Query Match      1.3%; Score 14; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      20  CARQKQCEGRYTSN 33
        11  CARQKQCEGRYTSN 24

RESULT 4
US-10-764-390-262
; Sequence 262, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 262
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-262

Query Match      1.3%; Score 14; DB 16; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      143  FLGKDWGLEMSEY 156
Db      1  FLGKDWGLEMSEY 14

RESULT 5
US-10-764-390-268
; Sequence 268, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 268
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-268
```

```
Query Match      1.3%; Score 14; DB 16; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      128 GIWGDSPEDIRKDL 141
Db      1 GIWGDSPEDIRKDL 14

RESULT 6
US-10-764-390-44
; Sequence 44, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-44

Query Match      0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 TGVLSLLLL 13
Db      1 TGVLSLLLL 9

RESULT 7
US-10-764-390-45
; Sequence 45, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-45

Query Match      1.3%; Score 14; DB 16; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      128 GIWGDSPEDIRKDL 141
Db      1 GIWGDSPEDIRKDL 14

RESULT 6
US-10-764-390-44
; Sequence 44, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-44

Query Match      0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 TGVLSLLLL 13
Db      1 TGVLSLLLL 9

RESULT 7
US-10-764-390-45
; Sequence 45, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-45

Query Match      0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      22 RKQCSEGR 30
Db      1 RKQCSEGR 9

RESULT 9
US-10-764-390-47
; Sequence 47, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-47

Query Match      0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      22 RKQCSEGR 30
Db      1 RKQCSEGR 9

RESULT 9
US-10-764-390-47
; Sequence 47, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-47
```

```
; SEQ ID NO 47
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-47

Query Match          0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 GRTYSNAVI 36
   |||||
Db 1 GRTYSNAVI 9

RESULT 10
US-10-764-390-48
; Sequence 48, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254P1D6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-48

Query Match          0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 NAVISPNLE 41
   |||||
Db 1 NAVISPNLE 9

RESULT 11
US-10-764-390-49
; Sequence 49, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254P1D6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
```

```
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-49

Query Match          0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 IMRVSHTPP 53
   |||||
Db 1 IMRVSHTPP 9

RESULT 12
US-10-764-390-50
; Sequence 50, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254P1D6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-50

Query Match          0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 CCDLSSCDL 69
   |||||
Db 1 CCDLSSCDL 9

RESULT 13
US-10-764-390-51
; Sequence 51, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254P1D6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
```

; CURRENT APPLICATION NUMBER: US/10/764,390
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-51

Query Match 0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 CDLSGCCDLA 70
Db 1 CDLSGCCDLA 9
|||||

RESULT 14

US-10-764-390-52
; Sequence 52, Application US/10/764390
; Publication No. US20040214212A1

; GENERAL INFORMATION:

; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner

; APPLICANT: Juan J. Perez-Villar

; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins

; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of

; FILE REFERENCE: 51158-20081.00

; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24

; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-52

Query Match 0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 77 CYLVSCPHK 85
Db 1 CYLVSCPHK 9
|||||

RESULT 15

US-10-764-390-53

; Sequence 53, Application US/10/764390
; Publication No. US20040214212A1

; GENERAL INFORMATION:

; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner

; APPLICANT: Juan J. Perez-Villar

; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins

; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-53

Query Match 0.8%; Score 9; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 GPIRSYLTF 102
Db 1 GPIRSYLTF 9
|||||

Search completed: October 12, 2005, 10:27:51
Job time : 70 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:21:44 ; Search time 26 seconds
(without alignments)
3077.838 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 1072

Sequence: 1 MAPPTGVLSLLLVTTIAGC.....VSMNGSIRNGASFYCSKDR 1072

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 250370

Minimum DB seq length: 0

Maximum DB seq length: 35

Post-processing: Listing first 45 summaries

Database : 1: Issued Patents AA.*

2: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/pctus_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	0.6	7	2	US-08-666-473-53
2	6	0.6	10	3	US-09-001-984C-78
3	6	0.6	10	4	US-09-396-347F-78
4	6	0.6	11	3	US-09-261-855-14
5	6	0.6	11	4	US-09-873-637-14
6	6	0.6	12	2	US-08-096-762-205
7	6	0.6	12	3	US-09-042-353-320
8	6	0.6	12	3	US-08-758-417A-168
9	6	0.6	15	3	US-09-044-718-8
10	6	0.6	15	4	US-10-062-848-8
11	6	0.6	16	4	US-08-537-871A-41
12	6	0.6	16	4	US-08-537-871A-45
13	6	0.6	17	1	US-08-347-198A-11
14	6	0.6	17	2	US-08-248-839C-53
15	6	0.6	17	3	US-08-335-844A-35
16	6	0.6	17	3	US-08-335-844A-50
17	6	0.6	17	4	US-09-129-366-35
18	6	0.6	17	4	US-09-129-366-50
19	6	0.6	17	2	US-08-637-759B-185
20	6	0.6	18	2	US-08-280-864A-6
21	6	0.6	18	3	US-08-871-355A-185
22	6	0.6	18	3	US-09-092-291-6
23	6	0.6	18	3	US-09-201-945-185
24	6	0.6	19	4	US-09-355-040-27
25	6	0.6	20	2	US-08-053-451B-122
26	6	0.6	20	4	US-09-053-611-17
27	6	0.6	20	4	US-09-889-136-6

Sequence 11, Appl
Sequence 6, Appl
Sequence 15, Appl
Sequence 6, Appl
Sequence 4, Appl
Sequence 6, Appl
Sequence 14, Appl
Sequence 14, Appl
Sequence 617, App
Sequence 229, App
Sequence 232, App
Sequence 1290, Ap
Sequence 13, Appl
Sequence 10, Appl
Sequence 3, Appl
Sequence 4, Appl

28 6 0.6 20 4 US-09-889-136-11
29 6 0.6 22 1 US-08-483-146A-6
30 6 0.6 22 1 US-08-232-513A-15
31 6 0.6 22 1 US-08-484-594A-6
32 6 0.6 22 3 US-09-231-159-4
33 6 0.6 22 3 US-08-611-307-4
34 6 0.6 22 4 US-09-076-258A-6
35 6 0.6 24 1 US-08-628-291-14
36 6 0.6 24 2 US-09-128-722-14
37 6 0.6 25 4 US-09-205-258-617
38 6 0.6 26 3 US-08-467-023-229
39 6 0.6 26 3 US-08-467-023-232
40 6 0.6 27 4 US-09-471-276-1290
41 6 0.6 27 4 US-09-936-885A-13
42 6 0.6 28 4 US-09-181-341-10
43 6 0.6 29 3 US-09-348-578-4
44 6 0.6 29 3 US-09-023-731-3
45 6 0.6 29 4 US-09-699-684-4

ALIGNMENTS

RESULT 1

US-08-666-473-53
; Sequence 53, Application US/086666473
; Patent No. 5843713
; GENERAL INFORMATION:
; APPLICANT: YOSHIDA, Aruto
; APPLICANT: TAKEUCHI, Makoto
; TITLE OF INVENTION: PEPTIDE SEQUENCE THAT FORMS MUCIN SUGAR
; TITLE OF INVENTION: CHAIN AND TECHNIQUE FOR MODIFYING PROTEIN TO BE LINKED
; TITLE OF INVENTION: WITH MUCIN SUGAR CHAIN
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,473
; FILING DATE: 19-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP95/02238
; FILING DATE: 01-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-22101
; FILING DATE: 09-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-269111
; FILING DATE: 01-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/837
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-666-473-53

Query Match 0.6%; Score 6; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 222 ASTPAP 227
|||||
Db 2 ASTPAP 7

RESULT 2

US-09-001-984C-78
; Sequence 78, Application US/09001984C
; Patent No. 6245331
; GENERAL INFORMATION:
; APPLICANT: Laal, Suman
; APPLICANT: Zolla-Pazner, Susan
; APPLICANT: Belisle, John T
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
; FILE REFERENCE: NYU-011
; CURRENT APPLICATION NUMBER: US/09/001,984C
; CURRENT FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: 60/034,003
; PRIOR FILING DATE: 1996-12-31
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis strain H37Rv
US-09-001-984C-78

Query Match 0.6%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 APAPPV 364
|||||
Db 3 APAPPV 8

RESULT 3

US-09-396-347F-78
; Sequence 78, Application US/09396347F
; Patent No. 6506384
; GENERAL INFORMATION:
; APPLICANT: Laal, Suman
; APPLICANT: Zolla-Pazner, Susan
; APPLICANT: Belisle, John T
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE
; FILE REFERENCE: 32004-189276
; CURRENT APPLICATION NUMBER: US/09/396,347F
; CURRENT FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: 09/001,984
; PRIOR FILING DATE: 1997-12-31
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis strain H37Rv
US-09-396-347F-78

Query Match 0.6%; Score 6; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 APAPPV 364
|||||
Db 3 APAPPV 8

RESULT 4

US-09-261-855-14
; Sequence 14, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-14

Query Match 0.6%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 442 LQELTL 447
|||||
Db 5 LQELTL 10

RESULT 5

US-09-873-637-14
; Sequence 14, Application US/09873637
; Patent No. 6794151
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-637-14

Query Match 0.6%; Score 6; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 442 LQELTL 447
|||||
Db 5 LQELTL 10

RESULT 6

US-08-096-762-205
; Sequence 205, Application US/08096762
; Patent No. 5814318
; GENERAL INFORMATION:
; APPLICANT: Lonberg, Nils
; APPLICANT: Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 5814318-Human Animals for
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 200
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,762
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,131
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; APPLICATION NUMBER: US 07/904,068
; FILING DATE: 23-JUN-1992
; APPLICATION NUMBER: US 07/853,408
; FILING DATE: 18-MAR-1992
; APPLICATION NUMBER: US 07/810,279
; FILING DATE: 17-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William W.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14643-9-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 205:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-096-762-205

Query Match 0.6%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 VTVSSE 412
Db 7 VTVSSE 12

RESULT 7
US-09-042-353-320
; Sequence 320, Application US/09042353
; Patent No 6255458
; GENERAL INFORMATION:
; APPLICANT: Lonberg, Nils
; APPLICANT: Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
; TITLE OF INVENTION: Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 421
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/042,353
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/810,279
; FILING DATE: 17-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/853,408
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/904,068
; FILING DATE: 23-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; APPLICATION NUMBER: US 08/053,131
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/096,762
; FILING DATE: 22-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/155,301
; FILING DATE: 18-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,739
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/165,699
; FILING DATE: 10-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,741
; FILING DATE: 09-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/352,322
; FILING DATE: 07-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/544,404
; FILING DATE: 10-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/728,463
; FILING DATE: 10-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US96/16433
; FILING DATE: 10-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,417
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/21803
; FILING DATE: 01-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 014643-00904005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 320:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-042-353-320

Query Match 0.6%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 VTVSSE 412
Db 7 VTVSSE 12

```

Db 7 VTVSSE 12

RESULT 8
US-08-758-417A-168
; Sequence 168, Application US/08758417A
; Patent No. 6300129
; GENERAL INFORMATION:
; APPLICANT: Lonberg, Nils
; Kay, Robert M.
; TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for
; Producing Heterologous Antibodies
; NUMBER OF SEQUENCES: 417
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758,417A
; FILING DATE: 02-Dec-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/728,463
; FILING DATE: 10-OCT-1996
; APPLICATION NUMBER: US 08/544,404
; FILING DATE: 10-OCT-1995
; APPLICATION NUMBER: US 08/352,322
; FILING DATE: 07-DEC-1994
; APPLICATION NUMBER: US 08/209,741
; FILING DATE: 09-MAR-1994
; APPLICATION NUMBER: US 08/165,699
; FILING DATE: 10-DEC-1993
; APPLICATION NUMBER: US 08/161,739
; FILING DATE: 03-DEC-1993
; APPLICATION NUMBER: US 08/155,301
; FILING DATE: 18-NOV-1993
; APPLICATION NUMBER: US 08/096,762
; FILING DATE: 22-JUL-1993
; APPLICATION NUMBER: US 08/053,131
; FILING DATE: 26-APR-1993
; APPLICATION NUMBER: US 07/990,860
; FILING DATE: 16-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Serafini, Andrew T.
; REGISTRATION NUMBER: 41,303
; REFERENCE/DOCKET NUMBER: 014643-009030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-08-758-417A-168

Query Match 0.6%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 VTVSSE 412
|||||

Db 7 VTVSSE 12

RESULT 9
US-09-044-718-8
; Sequence 8, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Talaromyces thermophilus
US-09-044-718-8

Query Match 0.6%; Score 6; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SLLLLV 15
|||||
Db 2 SLLLLV 7

RESULT 10
US-10-062-848-8
; Sequence 8, Application US/10062848
; Patent No. 6734004
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/10/062,848
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/044,718
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Talaromyces thermophilus
US-10-062-848-8

Query Match 0.6%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SLLLLV 15
|||||
Db 2 SLLLLV 7

```
RESULT 11
US-08-537-871A-41
; Sequence 41, Application US/08537871A
; Patent No. 6838254
; GENERAL INFORMATION:
; APPLICANT: Raymond HAMERS
; APPLICANT: Cecile HAMERS-CASERMAN
; APPLICANT: Serge V. M. MUYLDERMANS
; APPLICANT: Leon G. J. FRENKEN
; APPLICANT: Cornelis T. VERRIPS
; TITLE OF INVENTION: Production of antibodies or (functionalized)
; fragments thereof derived from heavy chain immunoglobulins
; TITLE OF INVENTION: of Camelidae
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,871A
; FILING DATE: 29-JAN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/01442
; FILING DATE: 28-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93202079.5
; FILING DATE: 15-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93201454.1
; FILING DATE: 19-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93201239.6
; FILING DATE: 29-APR-1993
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: See figure 6
;
US-08-537-871A-41

Query Match 0.6%; Score 6; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 VTSVSE 412
Db 1 VTSVSE 6
|||||

RESULT 12
US-08-537-871A-45
; Sequence 45, Application US/08537871A
; Patent No. 6838254
; GENERAL INFORMATION:
; APPLICANT: Raymond HAMERS
; APPLICANT: Cecile HAMERS-CASERMAN
; APPLICANT: Serge V. M. MUYLDERMANS
; APPLICANT: Leon G. J. FRENKEN
; APPLICANT: Cornelis T. VERRIPS
; TITLE OF INVENTION: Production of antibodies or (functionalized)
; fragments thereof derived from heavy chain immunoglobulins
; TITLE OF INVENTION: of Camelidae
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,871A
; FILING DATE: 29-JAN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/01442
; FILING DATE: 28-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93202079.5
; FILING DATE: 15-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93201454.1
; FILING DATE: 19-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93201239.6
; FILING DATE: 29-APR-1993
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: See figure 6
;
US-08-537-871A-41

Query Match 0.6%; Score 6; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 VTSVSE 412
Db 1 VTSVSE 6
|||||

RESULT 13
US-08-347-198A-11
; Sequence 11, Application US/08347198A
; Patent No. 5747046
; GENERAL INFORMATION:
; APPLICANT: MUNN, Edward A.
; APPLICANT: SMITH, Trevor S.
; TITLE OF INVENTION: PRODUCTION AND USE OF ANTHELMINTIC
; AGENTS AND PROTECTIVE IMMUNOGENS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W., Suite 330 - G
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; TITLE OF INVENTION: of Camelidae
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,871A
; FILING DATE: 29-JAN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/01442
; FILING DATE: 28-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93202079.5
; FILING DATE: 15-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93201454.1
; FILING DATE: 19-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 93201239.6
; FILING DATE: 29-APR-1993
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: See figure 19
;
US-08-537-871A-45

Query Match 0.6%; Score 6; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 VTSVSE 412
Db 1 VTSVSE 6
|||||

RESULT 13
US-08-347-198A-11
; Sequence 11, Application US/08347198A
; Patent No. 5747046
; GENERAL INFORMATION:
; APPLICANT: MUNN, Edward A.
; APPLICANT: SMITH, Trevor S.
; TITLE OF INVENTION: PRODUCTION AND USE OF ANTHELMINTIC
; AGENTS AND PROTECTIVE IMMUNOGENS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W., Suite 330 - G
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/08/347,198A
 ; FILING DATE: 21-NOV-1994
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/020,526
 ; FILING DATE: 22-FEB-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/761,749
 ; FILING DATE: 17-SEP-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB90/00416
 ; FILING DATE: 19-MAR-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 89906156.8
 ; FILING DATE: 17-MAR-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KITTs, Monica C.
 ; REGISTRATION NUMBER: 36,105
 ; REFERENCE/DOCKET NUMBER: P443-1289
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202/638-5000
 ; TELEFAX: 202/638-4810
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 17 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-347-198A-11

Query Match 0.6%; Score 6; DB 1; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 881 KAAEVA 886
 Db 1 KAAEVA 6

RESULT 14
 US-08-248-839C-53
 ; Sequence 53, Application US/08248839C
 ; Patent No. 5843702
 ; GENERAL INFORMATION:
 ; APPLICANT: McConnell, David
 ; APPLICANT: Devine, Kevin
 ; APPLICANT: O'Kane, Charles
 ; TITLE OF INVENTION: A Gene Expression System
 ; NUMBER OF SEQUENCES: 185
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 5843702o No. 5843702disk of No. 5843702th America, Inc.
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/248,839C
 ; FILING DATE: 25-MAY-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gregg, Valeta A.
 ; REGISTRATION NUMBER: 35,127
 ; REFERENCE/DOCKET NUMBER: 3614.214-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655

; INFORMATION FOR SEQ ID NO: 53:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 17 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Protein
 ; US-08-248-839C-53

Query Match 0.6%; Score 6; DB 2; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 LTSALI 454
 Db 5 LTSALI 10

RESULT 15
 US-08-335-844A-35
 ; Sequence 35, Application US/08335844A
 ; Patent No. 6066503

; GENERAL INFORMATION:
 ; APPLICANT: GRAHAM, MARGARET
 ; APPLICANT: SMITH, TREVOR STANLEY
 ; APPLICANT: MUNN, EDWARD ALBERT
 ; APPLICANT: KNOX, DAVID PATRICK
 ; APPLICANT: OLIVER, JOANNA JANE
 ; APPLICANT: NEWTON, SUSAN ELIZABETH
 ; TITLE OF INVENTION: RECOMBINANT DNA MOLECULES ENCODING
 ; TITLE OF INVENTION: AMINOPEPTIDASE ENZYMES AND THEIR USE IN THE PREPARATION OF
 ; TITLE OF INVENTION: VACCINES AGAINST HELMINTH INFECTIONS
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rothwell, Figg, Ernst & Kurz
 ; STREET: Suite 701-E, 555 Thirteenth St., N.W
 ; CITY: Washington
 ; STATE: D. C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20004

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/335,844A
 ; FILING DATE: 09-JAN-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB PCT/GB93/00943
 ; FILING DATE: 06-MAY-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9209936
 ; FILING DATE: 08-MAY-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WALKER, Barbara W.
 ; REGISTRATION NUMBER: 35,400
 ; REFERENCE/DOCKET NUMBER: 1181-223A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)783-6040
 ; TELEFAX: (202)783-6031
 ; INFORMATION FOR SEQ ID NO: 35:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 17 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-335-844A-35

Query Match 0.6%; Score 6; DB 3; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 881 KAAEVA 886

Db 1 KAAEVA 6

Search completed: October 12, 2005, 10:26:31
Job time : 27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:18:15 ; Search time 68 Seconds

(without alignments)
8072.778 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 1072

Sequence: 1 MAPPTGVLSLLLVTIAGC.....VSMNGSRNGASFYCSKDR 1072

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 37082

Minimum DB seq length: 0

Maximum DB seq length: 35

Post-processing: Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	6	0.6	18	2 Q71UR5	Q71ur5 homo sapien
2	6	0.6	22	2 Q924C7	Q924c7 mus musculus
3	6	0.6	24	2 Q8HWR5	Q8hwr5 homo sapien
4	6	0.6	24	2 Q8HWR6	Q8hwr6 homo sapien
5	6	0.6	24	2 Q8HWR7	Q8hwr7 homo sapien
6	6	0.6	24	2 Q8HWR9	Q8hwr9 homo sapien
7	6	0.6	24	2 Q8HWS0	Q8hws0 homo sapien
8	6	0.6	24	2 Q8HWS1	Q8hws1 homo sapien
9	6	0.6	24	2 Q8HWS2	Q8hws2 homo sapien
10	6	0.6	24	2 Q8MGZ0	Q8mgz0 homo sapien
11	6	0.6	24	2 Q861D7	Q861d7 homo sapien
12	6	0.6	26	1 HCY3_HOMAM	P82298 homarus ame
13	6	0.6	27	1 DMS4_PHYSA	P80280 phyllomedus
14	6	0.6	28	2 Q86618	Q86618 human herpe
15	6	0.6	31	2 Q71VQ3	Q71vq3 oryctolagus
16	6	0.6	32	2 Q7M4E6	Q7m4e6 drosophila
17	6	0.6	33	2 Q52315	Q52315 escherichia
18	6	0.6	35	1 HCYA_CHEDE	P83173 cherax dest
19	6	0.6	35	2 Q81WM8	Q81wm8 homo sapien
20	6	0.6	35	2 Q80595	Q80595 human immun
21	5	0.5	9	1 PPH1_LYCES	P83380 lycopersico
22	5	0.5	9	2 Q6UIQ3	Q6uiq3 macaca mula
23	5	0.5	9	2 Q70Y80	Q70y80 plectranthu
24	5	0.5	9	2 Q6Q7F3	Q6q7f3 alcaligenes
25	5	0.5	10	2 Q7S225	Q7s225 neurospora
26	5	0.5	10	2 Q68SM8	Q68sm8 chaetodipus
27	5	0.5	10	2 Q68SM9	Q68sm9 thomomys mo
28	5	0.5	10	2 Q68SNO	Q68sno cratogeomys
29	5	0.5	10	2 Q68SN1	Q68sn1 cratogeomys
30	5	0.5	10	2 Q6LDC1	Q6ldc1 mus sp. pla
31	5	0.5	10	2 Q39949	Q39949 gb virus c/

32	5	0.5	10	2	Q39957	Q39957 gb virus c/
33	5	0.5	10	2	Q39958	Q39958 gb virus c/
34	5	0.5	10	2	Q77VW7	Q77vw7 gb virus c/
35	5	0.5	10	2	Q77VW8	Q77vw8 gb virus c/
36	5	0.5	10	2	Q77VJ1	Q77vj1 gb virus c/
37	5	0.5	10	2	Q77ZJ2	Q77zj2 gb virus c/
38	5	0.5	10	2	Q77ZJ3	Q77zj3 gb virus c/
39	5	0.5	10	2	Q77ZJ4	Q77zj4 gb virus c/
40	5	0.5	10	2	Q9W910	Q9w910 gb virus c/
41	5	0.5	10	2	Q71HJ9	Q71hj9 icterus spu
42	5	0.5	11	2	Q47569	Q47569 escherichia
43	5	0.5	12	2	Q9X6V0	Q9x6v0 aquifex pyr
44	5	0.5	13	2	Q818F2	Q818f2 drosophila
45	5	0.5	13	2	Q7M355	Q7m355 ovis aries

ALIGNMENTS

RESULT 1

Q71UR5	Q71UR5	PRELIMINARY;	PRT;	18 AA.
AC	Q71UR5;			
DT	05-JUL-2004 (TrEMBLrel. 27, Created)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	Norepinephrine transporter (Fragment).			
GN	Name=NET;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99156934; PubMed=10037744; DOI=10.1074/jbc.274.10.6507;			
RA	Kim C.H., Kim H.S., Cubellis J.F., Kim K.S.;			
RT	"A previously undescribed intron and extensive 5' upstream sequence,			
RT	but not Phox2a-mediated transactivation, are necessary for high level			
RT	cell type-specific expression of the human norepinephrine transporter			
RT	gene.";			
RL	J. Biol. Chem. 274:6507-6518 (1999).			
DR	EMBL; AF061198; AAD17521.1; -.			
FT	NON TER	18	18	
SQ	SEQUENCE	18 AA;	1998 MW;	E8333908ED35CD2D CRC64;

Query Match 0.6%; Score 6; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	617	VQPENN	622
Db	10	VQPENN	15

RESULT 2

Q924C7	Q924C7	PRELIMINARY;	PRT;	22 AA.
ID	Q924C7;			
AC	Q924C7;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Glucagon-like peptide-2 receptor (Fragment).			
GN	Name=Glp2r;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/SVJ;			
RC	MEDLINE=21292988; PubMed=11262390; DOI=10.1074/jbc.M009382200;			
RX	Loveshin J.A., Estall J., Yusta B., Brown T.J., Drucker D.J.;			
RA	"Glucagon-like peptide (GLP)-2 action in the murine central nervous			
RT				

RT system is enhanced by elimination of GLP-1 receptor signaling.";

RL J. Biol. Chem. 276:21489-21499(2001).

DR EMBL; AF338224; AAK63043.1; --.

DR MGD; MGI:2136733; Glp2r.

DR GO; GO:0016021; C:integral to membrane; TAS.

DR GO; GO:0004967; F:glucagon receptor activity; TAS.

KW GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.

KW Receptor.

FT NON_TER 22 22

SQ SEQUENCE 22 AA; 2526 MW; 2C5BF53DCCD425C9 CRC64;

Query Match 0.6%; Score 6; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 SLLLLV 15

|||||

Db 13 SLLLLV 18

RESULT 3

Q8HWR5 PRELIMINARY; PRT; 24 AA.

ID Q8HWR5;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Leucocyte antigen (Fragment).

GN Name=HLA-B;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ316292; CAC87137.1; --.

FT NON_TER 24 24

SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;

Best Local Similarity 100.0%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332

|||||

Db 4 TAPRTV 9

RESULT 4

Q8HWR6 PRELIMINARY; PRT; 24 AA.

ID Q8HWR6;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Leucocyte antigen (Fragment).

GN Name=HLA-B;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ316291; CAC87136.1; --.

FT NON_TER 24 24

SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;

Best Local Similarity 100.0%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332

|||||

Db 4 TAPRTV 9

RESULT 7

Q8HWS0 PRELIMINARY; PRT; 24 AA.

ID Q8HWS0;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332

|||||

Db 4 TAPRTV 9

RESULT 5

Q8HWR7 PRELIMINARY; PRT; 24 AA.

ID Q8HWR7;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Leucocyte antigen (Fragment).

GN Name=HLA-B;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ316290; CAC87135.1; --.

FT NON_TER 24 24

SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;

Best Local Similarity 100.0%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332

|||||

Db 4 TAPRTV 9

RESULT 6

Q8HWR9 PRELIMINARY; PRT; 24 AA.

ID Q8HWR9;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Leucocyte antigen (Fragment).

GN Name=HLA-B;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ316288; CAC87133.1; --.

FT NON_TER 24 24

SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;

Best Local Similarity 100.0%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332

|||||

Db 4 TAPRTV 9

```
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Leucocyte antigen (Fragment).
GN Name=HLA-B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ316287; CAC87132.1; -.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332
Db 4 TAPRTV 9

RESULT 8
Q8HWS1 ID Q8HWS1 PRELIMINARY; PRT; 24 AA.
AC Q8HWS1;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Leucocyte antigen (Fragment).
GN Name=HLA-B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ316286; CAC87879.1; -.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332
Db 4 TAPRTV 9

RESULT 9
Q8HWS2 ID Q8HWS2 PRELIMINARY; PRT; 24 AA.
AC Q8HWS2;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Leucocyte antigen (Fragment).
GN Name=HLA-B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ316285; CAC87878.1; -.
FT NON_TER 24
```

```
SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332
Db 4 TAPRTV 9

RESULT 10
Q8MGZ0 ID Q8MGZ0 PRELIMINARY; PRT; 24 AA.
AC Q8MGZ0;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Leucocyte antigen (Fragment).
GN Name=HLA-B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ429190; CAD22131.1; -.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332
Db 4 TAPRTV 9

RESULT 11
Q861D7 ID Q861D7 PRELIMINARY; PRT; 24 AA.
AC Q861D7;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE MHC class I antigen (Fragment).
GN Name=HLA-B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Peripheral blood;
RA Swelsen W.T., Voorter C.E., van den Berg-Loonen E.M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ417674; CAD10406.1; -.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2669 MW; 1942CFC2489C177F CRC64;

Query Match 0.6%; Score 6; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 TAPRTV 332
Db 4 TAPRTV 9

RESULT 12
```

```

HCY3 HOMAM
ID HCY3 HOMAM STANDARD; PRT; 26 AA.
AC P82298;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hemocyanin subunit 3 (Fragment).
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Nephropoidea; Nephropidae; Homarus.
NCBI_TaxID=6706;
RN [1]
SEQUENCE.
RP MEDLINE=99259579; PubMed=10327595; DOI=10.1016/S0305-0491(98)10144-X;
RA Stoeva S., Dolashka P., Hristova R., Genov N., Voelter W.;
RT "Subunit composition and N-terminal analysis of arthropod
hemocyanins."
RL Comp. Biochem. Physiol. 122B:69-75(1999).
CC -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers
occurring freely dissolved in the hemolymph of many mollusks and
arthropods.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Hemolymph.
CC -!- SIMILARITY: Belongs to the tyrosinase family. Hemocyanin
subfamily.
DR InterPro: IPR000896; Hemocyanin.
DR InterPro: IPR002227; Tyrosinase.
DR PROSITE; PS00209; HEMOCYANIN 1; PARTIAL.
DR PROSITE; PS00210; HEMOCYANIN 2; PARTIAL.
DR PROSITE; PS00497; TYROSINASE 1; PARTIAL.
DR PROSITE; PS00498; TYROSINASE 2; PARTIAL.
KW Copper; Direct protein sequencing; Hemolymph; Oxygen transport;
KW Transport.
FT NON_TER 26 26
SQ SEQUENCE 26 AA; 2903 MW; 7F3549875A638D61 CRC64;

Query Match 0.6%; Score 6; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 FLFLFKV 905
Db 19 FLFLFKV 24

RESULT 13
DMS4_PHYSA STANDARD; PRT; 27 AA.
AC P80280;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Dermaseptin 4 (DS IV).
OS Phyllomedusa sauvagei (Savage's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae;
OC Phyllomedusinae; Phyllomedusa.
OC NCBI_TaxID=8395;
RN [1]
SEQUENCE.
RP TISSUE=Skin secretion;
RX MEDLINE=94139686; PubMed=83065981;
RA Mor A., Nicolas P.;
RT "Isolation and structure of novel defensive peptides from frog skin."
RL Eur. J. Biochem. 219:145-154(1994).
CC -!- FUNCTION: Possesses a potent antimicrobial activity against
bacteria, fungi and protozoa. Probably acts by disturbing membrane
functions with its amphipathic structure.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Belongs to the frog skin active peptide (FSAP) family.
Dermaseptin subfamily.

```

```

KW Amphibian defense peptide; Antibiotic; Direct protein sequencing;
KW Fungicide; Multigene family.
SQ SEQUENCE 27 AA; 2779 MW; 43C94D2DC19721A8 CRC64;

Query Match 0.6%; Score 6; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 878 KVLKAA 883
Db 9 KVLKAA 14

RESULT 14
Q86618 PRELIMINARY; PRT; 28 AA.
ID Q86618;
AC Q86618;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE UL15 protein (Fragment).
OS Name=UL15;
OS Human herpesvirus 1 (HHV-1) (Human herpes simplex virus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OC NCBI_TaxID=10298;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=92024477; PubMed=1656627; DOI=10.1016/0168-1702(91)90064-3;
RX Dolan A., Aruckle M., McGeoch D.J.;
RT "Sequence analysis of the splice junction in the transcript of herpes
simplex virus type 1 gene UL15."
RL Virus Res. 20:97-104(1991).
DR EMBL; S62788; AAB20055.1; -.
DR GO; GO:0006323; P:DNA packaging; IEA.
DR InterPro; IPR003499; DNA_pack_N.
DR Pfam; PF02500; DNA_pack_N; 1.
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 3051 MW; CAE2EFDD17528458 CRC64;

Query Match 0.6%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 739 DGSRST 744
Db 2 DGSRST 7

RESULT 15
Q71VQ3 PRELIMINARY; PRT; 31 AA.
ID Q71VQ3;
AC Q71VQ3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Beta-actin (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RN [1]
SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand;
RA Carter E.P., Umenishi F., Matthay M.A., Verkman A.S.;
RT "Increased water permeability across the blood-gas barrier in rabbit
lungs in the first 24 hours after birth."
RL J. Clin. Invest. 0:0-0(1997).
DR EMBL; AF000313; AAB94410.1; -.
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 3525 MW; F80D65C92838FA92 CRC64;

```

Query Match 0.6%; Score 6; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 DIRKDL 141
|||
Db 24 DIRKDL 29

Search completed: October 12, 2005, 10:25:21
Job time : 69 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:19:24 ; Search time 21 Seconds
(without alignments)
4911.637 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 1072
Sequence: 1 MAPPTGVLSLLLLVVIAGC.....VSMNGSIRNGASFYCSKDR 1072

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7756

Minimum DB seq length: 0
Maximum DB seq length: 35

Post-processing: Listing first 45 summaries

Database : PIR 79:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	0.6	32	2 S55937	translation initia
2	6	0.6	33	2 I53186	IS1/cat protein -
3	6	0.6	35	2 G60529	hemocyanin M3' - c
4	5	0.5	8	2 A21440	variant surface gl
5	5	0.5	10	2 A47384	placental lactogen
6	5	0.5	12	2 P70255	Ig heavy chain CRD
7	5	0.5	12	2 PT0319	Ig heavy chain CRD
8	5	0.5	13	2 A38929	glutathione peroxi
9	5	0.5	13	2 S09395	hypothetical prote
10	5	0.5	13	2 S47377	T-cell antigen rec
11	5	0.5	14	2 S29878	Na+/K+-exchanging
12	5	0.5	14	2 S27140	hypothetical prote
13	5	0.5	15	2 S26791	Ig heavy chain V r
14	5	0.5	15	2 P80382	Ig heavy chain J r
15	5	0.5	15	2 G41299	T-cell receptor al
16	5	0.5	15	2 P28587	T-cell receptor be
17	5	0.5	16	2 P80383	Ig heavy chain J r
18	5	0.5	16	2 D58501	26K kidney and gal
19	5	0.5	17	2 C37520	glutathione transf
20	5	0.5	17	2 S26747	Ig heavy chain J r
21	5	0.5	17	2 S26744	Ig heavy chain J r
22	5	0.5	17	2 I24687	T-cell receptor be
23	5	0.5	17	2 P80384	Ig heavy chain J r
24	5	0.5	17	2 A49237	45/47K antigen - M
25	5	0.5	18	2 S24780	protein-tyrosine k
26	5	0.5	18	2 A25941	Ig heavy chain J-H
27	5	0.5	18	4 I39461	anti-angiotensin,
28	5	0.5	19	2 S38837	T-cell receptor be
29	5	0.5	19	2 S28396	T-complex protein

30	5	0.5	19	2 PT0332	Ig heavy chain CRD
31	5	0.5	20	2 A60822	cytochrome P450 PB
32	5	0.5	20	2 S04961	malate dehydrogena
33	5	0.5	20	2 FC1151	equisetoxin 1C - s
34	5	0.5	20	2 B44920	2-halobenzoate 1,2
35	5	0.5	20	2 B44835	dTDPglucose 4,6-de
36	5	0.5	20	2 S63483	coenzyme P420 hydr
37	5	0.5	20	2 B34016	tenebrosin B - sea
38	5	0.5	20	2 B46174	RNA-binding protei
39	5	0.5	21	2 S69371	quodenase - bovine
40	5	0.5	21	2 A61487	110K ATPase, intes
41	5	0.5	21	2 S29858	ribosomal protein
42	5	0.5	21	2 I54268	alpha-1-antichymot
43	5	0.5	22	2 PH1333	Ig heavy chain DJ
44	5	0.5	22	2 D47256	kinetoplast DNA-as
45	5	0.5	23	2 A60529	hemocyanin 1 - edi

ALIGNMENTS

RESULT 1

S55937 translation initiation factor eIF-4E - fruit fly (Drosophila melanogaster) (fragments)
C:Species: Drosophila melanogaster
C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S55937
R:Hernandez, G.; Sierra, J.M.
Biochim. Biophys. Acta 1261, 427-431, 1995
A>Title: Translation initiation factor eIF-4E from Drosophila: cDNA sequence and express
A:Reference number: S55936; MUID:95260867; PMID:7742371
A:Accession: S55937
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8;9-27;28-32 <HER>
A:Cross-references: UNIPROT:Q7M4E6
C:Genetics:
A:Gene: FlyBase:EIF4F
A:Cross-references: FlyBase:FBgn0013947
C:Superfamily: translation initiation factor eIF-4E

Query Match 0.6%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.8e+02; Mismatches 0; Gaps 0;
Matches 6; Conservative 0; Indels 0;

Qy	307 PPTSAA 312
Db	21 PPTSAA 26

RESULT 2

I53186 IS1/cat protein - Escherichia coli (fragment)
C:Species: Escherichia coli
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C:Accession: I53186; I66859
R:Lida, S.; Marcoli, R.; Bickler, T.A.
EMBO J. 1, 755-759, 1982
A>Title: Phenotypic reversion of an IS1-mediated deletion mutation: A combined role for
A:Reference number: I53186; MUID:84236045; PMID:6329702
A:Accession: I53186
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-33 <RES>
A:Cross-references: UNIPROT:Q52315; GB:M24180; NID:g151758; PIDN:AAA26059.1; PID:g151759
A:Accession: I66859
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-18,'G' <RES2>
A:Cross-references: GB:M24181; NID:g151760; PIDN:AAA26060.1; PID:g151761

Query Match 0.6%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;


```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 867 IVFYVQ 872
Db 15 IVFYVQ 20

RESULT 3
G60529
hemocyanin M3' - crayfish (Cherax destructor) (fragment)
C:Species: Cherax destructor (Yabby)
C>Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 09-Jul-2004
C:Accession: G60529
R:Neuteboom, B.; Sierdsema, S.J.; Beintema, J.J.
Comp. Biochem. Physiol. B 94, 587-592, 1989
A>Title: The relationship between N-terminal sequences and immunological characterization
A:Reference number: A60529; MUID:90151075; PMID:2620501
A:Accession: G60529
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-35 <NEU>
A:Cross-references: UNIPROT:P83173

Query Match 0.6%; Score 6; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 900 FLLFKV 905
Db 17 FLLFKV 22

RESULT 4
A21440
variant surface glycoprotein pSLc1 - Trypanosoma brucei (fragment)
C:Species: Trypanosoma brucei
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jul-2004
C:Accession: A21440
R:Parsons, M.; Nelson, R.G.; Watkins, K.P.; Agabian, N.
Cell 38, 309-316, 1984
A>Title: Trypanosome mRNAs share a common 5' spliced leader sequence.
A:Reference number: A30853; MUID:84282716; PMID:6088073
A:Accession: A21440
A:Molecule type: mRNA
A:Residues: 1-8 <PAR>
A:Cross-references: UNIPROT:P22225; GB:K02195; NID:gl62150; PID:gl62151
C:Keywords: glycoprotein

Query Match 0.5%; Score 5; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 264 SGKEV 268
Db 2 SGKEV 6

RESULT 5
A47364
placental lactogen-I precursor - mouse (fragment)
C:Species: Mus sp. (mouse)
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-1996
C:Accession: A47364
R:Shida, M.M.; Ng, Y.K.; Soares, M.J.; Linzer, D.I.
Mol. Endocrinol. 7, 181-188, 1993
A>Title: Trophoblast-specific transcription from the mouse placental lactogen-I gene pro
A:Reference number: A47364; MUID:93225959; PMID:8469232
A:Accession: A47364
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-10 <RES>
A:Cross-references: GB:S58124; NID:Q299449
```

```
Query Match 0.5%; Score 5; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 TLNLS 396
Db 4 TLNLS 8

RESULT 6
PT0255
Ig heavy chain CRD3 region (clone 2-115B) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0255
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A>Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0255
A:Molecule type: DNA
A:Residues: 1-12 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 5; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 RVSHT 51
Db 3 RVSHT 7

RESULT 7
PT0319
Ig heavy chain CRD3 region (clone 6-127) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0319
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A>Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0319
A:Molecule type: DNA
A:Residues: 1-12 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 5; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 627 AVAGP 631
Db 6 AVAGP 10

RESULT 8
A38929
glutathione peroxidase (EC 1.11.1.9) - sheep (fragment)
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
C:Accession: A38929
R:Gottlieb, P.; Dyal, D.; Crews, B.
Arch. Biochem. Biophys. 294, 511-518, 1992
A>Title: Selenium-dependent glutathione peroxidases from ovine and bovine erythrocytes
A:Reference number: S21712; MUID:92231574; PMID:1567207
A:Accession: A38929
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-13 <GET>
```

A;Cross-references: UNIPROT:Q7M355
C;Superfamily: Glutathione peroxidase
C;Keywords: oxidoreductase

Query Match 0.5%; Score 5; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 APTV 332
Db 9 APTV 13

RESULT 9

S09395
hypothetical protein - fruit fly (*Drosophila melanogaster*) (fragment)
C;Species: *Drosophila melanogaster*
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 24-Jul-1997
C;Accession: S09395
R;Gisselmann, G.; Sewing, S.; Madseen, B.W.; Mallart, A.; Angaut-Petit, D.; Mueller-Holtk
EMBO J. 8, 2359-2364, 1989
A;Title: The interference of truncated with normal potassium channel subunits leads to a
A;Reference number: S09395; MUID:90005442; PMID:2551680
A;Accession: S09395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <GIS>

Query Match 0.5%; Score 5; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 260 SSNS 264
Db 9 SSNS 13

RESULT 10

S47377
T-cell antigen receptor VJ junction beta chain - human
C;Species: *Homo sapiens* (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: S47377
R;Lehner, P.J.
submitted to the EMBL Data Library, August 1994
A;Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T
A;Reference number: S47355
A;Accession: S47377
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LEH>
A;Cross-references: EMBL:Z35706; NID:9527505; PIDN:CAA84775.1; PID:9527506
C;Keywords: T-cell receptor

Query Match 0.5%; Score 5; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 741 SRSTD 745
Db 5 SRSTD 9

RESULT 11

S29878
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-2 chain - rat (fragment)
C;Species: *Rattus norvegicus* (Norway rat)
C;Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S29878
R;Lytton, J.
Biochem. Biophys. Res. Commun. 132, 764-769, 1995
A;Title: The catalytic subunits of the (Na⁺), K⁽⁺⁾-ATPase alpha and alpha(+) isozymes
A;Reference number: S29877; MUID:86050667; PMID:2998384

A;Accession: S29878
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <LVT>
A;Cross-references: UNIPROT:Q7M0G4
C;Keywords: hydrolase

Query Match 0.5%; Score 5; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 882 AAEVA 886
Db 7 AAEVA 11

RESULT 12

S27140
hypothetical protein 1 estrogen receptor 5'-region - human
C;Species: *Homo sapiens* (man)
C;Date: 05-Mar-1994 #sequence_revision 12-May-1995 #text_change 16-Aug-2004
C;Accession: S27140
R;Keaveney, M.; Klug, J.; Gammon, F.
DNA Seq. 2, 347-358, 1992
A;Title: Sequence analysis of the 5' flanking region of the human estrogen receptor gene
A;Reference number: S27140; MUID:93075998; PMID:1476547
A;Accession: S27140
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-14 <KEA>
A;Cross-references: EMBL:X62462; NID:g31201; PIDN:CAA44319.1; PID:g31202

Query Match 0.5%; Score 5; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 271 PSHSL 275
Db 8 PSHSL 12

RESULT 13

S26791
Ig heavy chain V region (N63P2) - human (fragment)
C;Species: *Homo sapiens* (man)
C;Date: 13-Jan-1995 #sequence_revision 02-Aug-1996 #text_change 20-Jun-2000
C;Accession: S26791; S19879
R;Mortari, F.; Newton, J.A.; Wang, J.Y.; Schroeder Jr., H.W.
Eur. J. Immunol. 22, 241-245, 1992
A;Title: The human cord blood antibody repertoire. Frequent usage of the V(H)7 gene fami
A;Reference number: S26786; MUID:92111632; PMID:1730251
A;Accession: S26791
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <MOR>
A;Cross-references: EMBL:X61022; NID:g32791; PIDN:CAA43356.1; PID:g1335124; EMBL:X61234;
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin

Query Match 0.5%; Score 5; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.6e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 VTVSS 411
Db 11 VTVSS 15

RESULT 14

PS0382
Ig heavy chain J region 2 - rat (fragment)
C;Species: *Rattus norvegicus* (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 06-Jun-1997

C;Accession: PS0382
R;Lang, P.; Mocikat, R.
Gene 102, 261-264, 1991
A;Title: Immunoglobulin heavy-chain joining genes in the rat: comparison with mouse and
A;Reference number: JH0666; MUID:91340162; PMID:1908401
A;Accession: PS0382
A;Molecule type: DNA
A;Residues: 1-15 <LAN>
A;Cross-references: EMBL:X56791
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 0.5%; Score 5; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 VTVSS 411
|||||
Db 11 VTVSS 15

RESULT 15

G41299
T-cell receptor alpha chain precursor J region (39) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 05-Nov-1999
C;Accession: G41299
R;Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.; Panayi, G.;
Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991
A;Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheumatoid
A;Reference number: A41299; MUID:92020887; PMID:1656449
A;Accession: G41299
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <UEM>
A;Cross-references: GB:S57457; NID:g236330; PIDN:AA19962.1; PID:g236331
C;Keywords: T-cell receptor

Query Match 0.5%; Score 5; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 RLTVK 696
|||||
Db 10 RLTVK 14

Search completed: October 12, 2005, 10:25:54
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:02:54 ; Search time 73 Seconds
(without alignments)
5679.550 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 1072
Sequence: 1 MAPPTGVLSLLLVITAGC.....VSMNGSRNGASFYCSKDR 1072

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 854512

Minimum DB seq length: 0
Maximum DB seq length: 35

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	1.7	19	4	ABG22066 Novel hum
2	16	1.5	16	8	ADR00632 Human 254
3	14	1.3	24	8	ADR00854 Human 254
4	14	1.3	29	8	ADR00851 Human 254
5	14	1.3	29	8	ADR00857 Human 254
6	9	0.8	9	8	ADR00673 Human 254
7	9	0.8	9	8	ADR00720 Human 254
8	9	0.8	9	8	ADR00722 Human 254
9	9	0.8	9	8	ADR00746 Human 254
10	9	0.8	9	8	ADR00759 Human 254
11	9	0.8	9	8	ADR00767 Human 254
12	9	0.8	9	8	ADR00797 Human 254
13	9	0.8	9	8	ADR00799 Human 254
14	9	0.8	9	8	ADR00835 Human 254
15	9	0.8	9	8	ADR00645 Human 254
16	9	0.8	9	8	ADR00666 Human 254
17	9	0.8	9	8	ADR00683 Human 254
18	9	0.8	9	8	ADR00693 Human 254
19	9	0.8	9	8	ADR00705 Human 254
20	9	0.8	9	8	ADR00706 Human 254
21	9	0.8	9	8	ADR00750 Human 254
22	9	0.8	9	8	ADR00756 Human 254
23	9	0.8	9	8	ADR00772 Human 254
24	9	0.8	9	8	ADR00775 Human 254
25	9	0.8	9	8	ADR00784 Human 254

26	9	0.8	9	8	ADR00790 Human 254
27	9	0.8	9	8	ADR00811 Human 254
28	9	0.8	9	8	ADR00814 Human 254
29	9	0.8	9	8	ADR00844 Human 254
30	9	0.8	9	8	ADR00639 Human 254
31	9	0.8	9	8	ADR00643 Human 254
32	9	0.8	9	8	ADR00648 Human 254
33	9	0.8	9	8	ADR00656 Human 254
34	9	0.8	9	8	ADR00671 Human 254
35	9	0.8	9	8	ADR00672 Human 254
36	9	0.8	9	8	ADR00680 Human 254
37	9	0.8	9	8	ADR00688 Human 254
38	9	0.8	9	8	ADR00689 Human 254
39	9	0.8	9	8	ADR00694 Human 254
40	9	0.8	9	8	ADR00721 Human 254
41	9	0.8	9	8	ADR00741 Human 254
42	9	0.8	9	8	ADR00754 Human 254
43	9	0.8	9	8	ADR00760 Human 254
44	9	0.8	9	8	ADR00782 Human 254
45	9	0.8	9	8	ADR00805 Human 254

ALIGNMENTS

RESULT 1
ABG22066
ID ABG22066 standard; protein; 19 AA.
XX
AC ABG22066;
XX
DT 18-FEB-2002 (first entry)
XX
DB Novel human diagnostic protein #22057.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS86253.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 52425; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

CC Sequence 19 AA;

Query Match 1.7%; Score 18; DB 4; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.7e-09; Mismatches 0; Indels 0; Gaps 0;
 Matches 18; Conservative 0;

QY 1 MAPPTGVLSSLLLVITIA 18
 |||||
 DB 1 MAPPTGVLSSLLLVITIA 18

RESULT 2
 ADR00632
 ID ADR00632 standard; peptide; 16 AA.
 XX
 AC ADR00632;

XX 04-NOV-2004 (first entry)

XX Human 254PID6B peptide SEQ ID NO:43.

XX 254PID6B; small interfering RNA; siRNA; immune response;
 KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
 KW chromosome 6.
 XX

XX Homo sapiens.

XX WO2004067716-A2.

XX 12-AUG-2004.

XX 23-JAN-2004; 2004WO-US001965.

XX 24-JAN-2003; 2003US-0442526P.

XX (AGEN-) AGENSYS INC.

XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Faris M;

XX WPI; 2004-580991/56.

XX New 254PID6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

XX Example 9; SEQ ID NO 43; 345pp; English.

XX The present invention describes a 254PID6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID
 CC NO:3.5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of PI; (2) a polynucleotide

CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of PI; (5)
 CC detecting, in a sample, the presence of a 254PID6B-related protein or a
 CC 254PID6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of PI; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of PI; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of PI. 254PID6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254PID6B peptide, which is used in the exemplification of the
 CC present invention. The human 254PID6B gene is located on chromosome 6p22.
 XX

CC Sequence 16 AA;

Query Match 1.5%; Score 16; DB 8; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07; Mismatches 0; Indels 0; Gaps 0;
 Matches 16; Conservative 0;

QY 522 NNAVDYPPVANAGPNH 537
 |||||
 DB 1 NNAVDYPPVANAGPNH 16

RESULT 3
 ADR00854

ID ADR00854 standard; peptide; 24 AA.

XX ADR00854;

XX 04-NOV-2004 (first entry)

XX Human 254PID6B v.3 peptide SEQ ID NO:265.

XX 254PID6B; small interfering RNA; siRNA; immune response;
 KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
 KW 254PID6B v.3; chromosome 6.
 XX

XX Homo sapiens.

XX WO2004067716-A2.

XX 12-AUG-2004.

XX 23-JAN-2004; 2004WO-US001965.

XX 24-JAN-2003; 2003US-0442526P.

XX (AGEN-) AGENSYS INC.

XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Faris M;

XX WPI; 2004-580991/56.

XX New 254PID6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

XX Example 13; SEQ ID NO 265; 345pp; English.

XX The present invention describes a 254PID6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,

consists essentially of, or consists of a peptide of eight, nine, ten, or eleven contiguous amino acids of a protein of figure 2 (p1, see SEQ ID NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in any of the 42 lists of peptides, given in the specification, or a protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or identical to an entire amino acid sequence of p1; (2) a polynucleotide that encodes the protein; (3) a composition comprising a polynucleotide that is a full complement of the polynucleotide described above; (4) generating a mammalian immune response directed to the protein of p1; (5) detecting, in a sample, the presence of a 254p1D6B-related protein or a 254p1D6B-related polynucleotide; (6) a composition that modulates the status of a cell that expresses a protein of p1; (7) a pharmaceutical composition that comprises the composition described above in a human unit dose form; (8) an antibody or its fragment, which is monoclonal; (9) a non-human transgenic animal that produces the antibody; (10) a hybridoma that produces the antibody; (11) delivering a cytotoxic agent or a diagnostic agent to a cell that expresses the protein of p1; and (12) inhibiting growth, reproduction or survival of cancer cells that express the protein of p1. 254p1D6B has cytostatic activity, and can be used in gene therapy. The compositions, molecules and methods are useful for treating and detecting cancer. The present sequence represents a human 254p1D6B v.3 peptide, which is used in the exemplification of the present invention. The human 254p1D6B gene is located on chromosome 6p22.

CC Sequence 24 AA;

Query Match 1.3%; Score 14; DB 8; Length 24;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 CARQCSGRTYSN 33
Db 11 CARQCSGRTYSN 24
|||||

RESULT 4
ADR00851
ID ADR00851 standard; peptide; 29 AA.
XX AC ADR00851;
XX DT 04-NOV-2004 (first entry)
XX DE Human 254p1D6B v.2 peptide SEQ ID NO:262.
XX KW 254p1D6B; small interfering RNA; siRNA; immune response;
XX KW 254p1D6B-related protein; cytostatic; gene therapy; cancer; human;
XX KW 254p1D6B v.2; chromosome 6.
XX OS Homo sapiens.
XX PN WO2004067716-A2.
XX PD 12-AUG-2004.
XX PF 23-JAN-2004; 2004WO-US001965.
XX PR 24-JAN-2003; 2003US-0442526P.
XX PA (AGEN-) AGENSYS INC.
XX PI Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
XX PI Perez-Villar JJ, Faris M;
XX DR WPI; 2004-580991/56.
XX PT New 254p1D6B siRNA composition comprising a double stranded siRNA that corresponds to the nucleic acid ORF sequence which encodes the 254p1D6B protein or corresponds to a subsequence of the ORF, useful for detecting and treating cancer.
XX PS Example 13; SEQ ID NO 262; 345pp; English.

The present invention describes a 254p1D6B small interfering RNA (siRNA) composition that comprises a double stranded siRNA that corresponds to the nucleic acid open reading frame (ORF) sequence which encodes the 254p1D6B protein, or corresponds to a subsequence of the ORF, where the double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides in length. Also described: (1) a composition that comprises, consists essentially of, or consists of a peptide of eight, nine, ten, or eleven contiguous amino acids of a protein of figure 2 (p1, see SEQ ID NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in any of the 42 lists of peptides, given in the specification, or a protein that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or identical to an entire amino acid sequence of p1; (2) a polynucleotide that encodes the protein; (3) a composition comprising a polynucleotide that is a full complement of the polynucleotide described above; (4) generating a mammalian immune response directed to the protein of p1; (5) detecting, in a sample, the presence of a 254p1D6B-related protein or a 254p1D6B-related polynucleotide; (6) a composition that modulates the status of a cell that expresses a protein of p1; (7) a pharmaceutical composition that comprises the composition described above in a human unit dose form; (8) an antibody or its fragment, which is monoclonal; (9) a non-human transgenic animal that produces the antibody; (10) a hybridoma that produces the antibody; (11) delivering a cytotoxic agent or a diagnostic agent to a cell that expresses the protein of p1; and (12) inhibiting growth, reproduction or survival of cancer cells that express the protein of p1. 254p1D6B has cytostatic activity, and can be used in gene therapy. The compositions, molecules and methods are useful for treating and detecting cancer. The present sequence represents a human 254p1D6B v.2 peptide, which is used in the exemplification of the present invention. The human 254p1D6B gene is located on chromosome 6p22.

CC Sequence 29 AA;

Query Match 1.3%; Score 14; DB 8; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 FLGKDWGLEMSEY 156
Db 1 FLGKDWGLEMSEY 14
|||||

RESULT 5
ADR00857
ID ADR00857 standard; peptide; 29 AA.
XX AC ADR00857;
XX DT 04-NOV-2004 (first entry)
XX DE Human 254p1D6B v.5 peptide SEQ ID NO:268.
XX DE 254p1D6B; small interfering RNA; siRNA; immune response;
XX KW 254p1D6B-related protein; cytostatic; gene therapy; cancer; human;
XX KW 254p1D6B v.5; chromosome 6.
XX OS Homo sapiens.
XX PN WO2004067716-A2.
XX PD 12-AUG-2004.
XX PF 23-JAN-2004; 2004WO-US001965.
XX PR 24-JAN-2003; 2003US-0442526P.
XX PA (AGEN-) AGENSYS INC.
XX PI Kanner SB, Raitano AB, Jakobovits A, Challita-Bid PM, Ge W;
XX PI Perez-Villar JJ, Faris M;
XX DR WPI; 2004-580991/56.
XX PT New 254p1D6B siRNA composition comprising a double stranded siRNA that

PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

XX Example 13; SEQ ID NO 268; 345pp; English.

XX The present invention describes a 254PID6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254PID6B-related protein or a
 CC 254PID6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254PID6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents a
 CC human 254PID6B v.5 peptide, which is used in the exemplification of the
 CC present invention. The human 254PID6B gene is located on chromosome 6p22.

XX Sequence 29 AA;

Query Match 1.3%; Score 14; DB 8; Length 29;

Best Local Similarity 100.0%; Pred. NO. 3.1e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 GIWGDSPEIDRKDL 141

Db 1 GIWGDSPEIDRKDL 14

RESULT 6

ID ADR00673

XX ADR00673 standard; peptide; 9 AA.

XX ADR00673;

XX 04-NOV-2004 (first entry)

XX Human 254PID6B peptide SEQ ID NO:84.

XX 254PID6B; small interfering RNA; siRNA; immune response;

KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
 KW chromosome 6.

XX Homo sapiens.

XX WO2004067716-A2.

XX 12-AUG-2004.

XX 23-JAN-2004; 2004WO-US0001965.

XX 24-JAN-2003; 2003US-0442526P.

XX (AGEN-) AGENSYS INC.

XX

PI Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 PI Perez-Villar JJ, Paris M;
 XX WPI; 2004-580991/56.

XX New 254PID6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

PS Example 9; SEQ ID NO 84; 345pp; English.

XX The present invention describes a 254PID6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254PID6B-related protein or a
 CC 254PID6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254PID6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254PID6B peptide, which is used in the exemplification of the
 CC present invention. The human 254PID6B gene is located on chromosome 6p22.

XX Sequence 9 AA;

Query Match 0.8%; Score 9; DB 8; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 320 ELPISPTTA 328

Db 1 ELPISPTTA 9

RESULT 7

ID ADR00720

XX ADR00720 standard; peptide; 9 AA.

XX ADR00720;

XX 04-NOV-2004 (first entry)

XX Human 254PID6B peptide SEQ ID NO:131.

XX 254PID6B; small interfering RNA; siRNA; immune response;

KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;
 KW chromosome 6.

XX Homo sapiens.

XX WO2004067716-A2.

XX 12-AUG-2004.

XX


```
DE Human 254P1D6B peptide SEQ ID NO:157.
XX
XX 254P1D6B; small interfering RNA; siRNA; immune response;
KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
XX chromosome 6.
XX Homo sapiens.
XX
XX WO2004067716-A2.
XX
XX 12-AUG-2004.
XX
XX 23-JAN-2004; 2004WO-US001965.
XX
XX 24-JAN-2003; 2003US-0442526P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
PI Perez-Villar JJ, Faris M;
XX
XX WPI; 2004-580991/56.
XX
XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX
XX Example 9; SEQ ID NO 157; 345pp; English.
XX
XX The present invention describes a 254P1D6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents the
CC human 254P1D6B peptide, which is used in the exemplification of the
CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
XX
XX Sequence 9 AA;
XX
XX Query Match 0.8%; Score 9; DB 8; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 47 RVSHTEFPV 55
XX |||||
XX 1 RVSHTEFPV 9
XX
XX RESULT 10
XX ADR00759
```

```
ID ADR00759 standard; peptide; 9 AA.
XX
XX ADR00759;
XX
XX 04-NOV-2004 (first entry)
XX
XX Human 254P1D6B peptide SEQ ID NO:170.
XX
XX 254P1D6B; small interfering RNA; siRNA; immune response;
KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
XX chromosome 6.
XX Homo sapiens.
XX OS
XX WO2004067716-A2.
XX
XX 12-AUG-2004.
XX
XX 23-JAN-2004; 2004WO-US001965.
XX
XX 24-JAN-2003; 2003US-0442526P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
PI Perez-Villar JJ, Faris M;
XX
XX WPI; 2004-580991/56.
XX
XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX
XX Example 9; SEQ ID NO 170; 345pp; English.
XX
XX The present invention describes a 254P1D6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
CC used in gene therapy. The compositions, molecules and methods are useful
CC for treating and detecting cancer. The present sequence represents the
CC human 254P1D6B peptide, which is used in the exemplification of the
CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
XX
XX Sequence 9 AA;
XX
XX Query Match 0.8%; Score 9; DB 8; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 305 PTPPTSAP 313
```

Db 1 PTPPTSAP 9
|||||
Query Match 0.8%; Score 9; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
ID ADR00767 standard; peptide; 9 AA.
XX ADR00767;
AC ADR00767;
XX 04-NOV-2004 (first entry)
XX Human 254P1D6B peptide SEQ ID NO:178.
XX 254P1D6B; small interfering RNA; siRNA; immune response;
KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
KW chromosome 6.
XX Homo sapiens.
XX WO2004067716-A2.
XX 12-AUG-2004.
XX 23-JAN-2004; 2004WO-US001965.
XX 24-JAN-2003; 2003US-0442526P.
XX (AGEN-) AGENSYS INC.
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
PI Perez-Villar JJ, Paris M;
XX WPI; 2004-580991/56.
XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX Example 9; SEQ ID NO 178; 345pp; English.

XX The present invention describes a 254P1D6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
XX Sequence 9 AA;

Query Match 0.8%; Score 9; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QV 363 PVEVTNYE 371
| | | | |
Db 1 PVEVTNYE 9

RESULT 12
ID ADR00797 standard; peptide; 9 AA.
XX ADR00797;
AC ADR00797;
XX 04-NOV-2004 (first entry)
XX Human 254P1D6B peptide SEQ ID NO:208.
XX 254P1D6B; small interfering RNA; siRNA; immune response;
KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
KW chromosome 6.
XX Homo sapiens.
XX WO2004067716-A2.
XX 12-AUG-2004.
XX 23-JAN-2004; 2004WO-US001965.
XX 24-JAN-2003; 2003US-0442526P.
XX (AGEN-) AGENSYS INC.
XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
PI Perez-Villar JJ, Paris M;
XX WPI; 2004-580991/56.
XX New 254P1D6B siRNA composition comprising a double stranded siRNA that
PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
PT protein or corresponds to a subsequence of the ORF, useful for detecting
PT and treating cancer.
XX Example 9; SEQ ID NO 208; 345pp; English.

XX The present invention describes a 254P1D6B small interfering RNA (siRNA)
CC composition that comprises a double stranded siRNA that corresponds to
CC the nucleic acid open reading frame (ORF) sequence which encodes the
CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
CC nucleotides in length. Also described: (1) a composition that comprises,
CC consists essentially of, or consists of a peptide of eight, nine, ten, or
CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
CC any of the 42 lists of peptides, given in the specification, or a protein
CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
CC that encodes the protein; (3) a composition comprising a polynucleotide
CC that is a full complement of the polynucleotide described above; (4)
CC generating a mammalian immune response directed to the protein of P1; (5)
CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
CC composition that comprises the composition described above in a human
CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
CC a non-human transgenic animal that produces the antibody; (10) a
CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
CC or a diagnostic agent to a cell that expresses the protein of P1; and
CC (12) inhibiting growth, reproduction or survival of cancer cells that
CC express the protein of P1. 254P1D6B has cytostatic activity, and can be

CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254PID6B peptide, which is used in the exemplification of the
 CC present invention. The human 254PID6B gene is located on chromosome 6p22.

XX Sequence 9 AA;

Query Match 0.8%; Score 9; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

QY 733 NNSITLDGS 741
 |||||

Db 1 NNSITLDGS 9

RESULT 13

ADRO0799

ID ADRO0799 standard; peptide; 9 AA.

XX ADRO0799;

DT 04-NOV-2004 (first entry)

DE Human 254PID6B peptide SEQ ID NO:210.

XX 254PID6B; small interfering RNA; siRNA; immune response;

KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;

KW chromosome 6.

XX Homo sapiens.

XX WO2004067716-A2.

XX 12-AUG-2004.

XX 23-JAN-2004; 2004WO-US001965.

XX 24-JAN-2003; 2003US-0442526P.

XX (AGEN-) AGENSYS INC.

XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;

PI Perez-Villar JJ, Faris M;

XX WPI; 2004-580991/56.

XX New 254PID6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

XX Example 9; SEQ ID NO 210; 345pp; English.

XX The present invention describes a 254PID6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of PI; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of PI; (5)
 CC detecting, in a sample, the presence of a 254PID6B-related protein or a
 CC 254PID6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of PI; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human

CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of PI; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of PI. 254PID6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254PID6B peptide, which is used in the exemplification of the
 CC present invention. The human 254PID6B gene is located on chromosome 6p22.

XX Sequence 9 AA;

Query Match 0.8%; Score 9; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 775 ALQLTNLVE 783

Db 1 ALQLTNLVE 9

RESULT 14

ADRO0835

ID ADRO0835 standard; peptide; 9 AA.

XX ADRO0835;

DT 04-NOV-2004 (first entry)

DE Human 254PID6B peptide SEQ ID NO:246.

XX 254PID6B; small interfering RNA; siRNA; immune response;

KW 254PID6B-related protein; cytostatic; gene therapy; cancer; human;

KW chromosome 6.

XX Homo sapiens.

XX WO2004067716-A2.

XX 12-AUG-2004.

XX 23-JAN-2004; 2004WO-US001965.

XX 24-JAN-2003; 2003US-0442526P.

XX (AGEN-) AGENSYS INC.

XX Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;

PI Perez-Villar JJ, Faris M;

XX WPI; 2004-580991/56.

XX New 254PID6B siRNA composition comprising a double stranded siRNA that
 PT corresponds to the nucleic acid ORF sequence which encodes the 254PID6B
 PT protein or corresponds to a subsequence of the ORF, useful for detecting
 PT and treating cancer.

XX Example 9; SEQ ID NO 246; 345pp; English.

XX The present invention describes a 254PID6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254PID6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or
 CC eleven contiguous amino acids of a protein of figure 2 (PI, see SEQ ID
 CC NO:3, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of PI; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide

CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254P1D6B peptide, which is used in the exemplification of the
 CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX

Sequence 9 AA;

Query Match 0.8%; Score 9; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 558 QIVLYEWSL 566
 Db 1 QIVLYEWSL 9
 |||||

RESULT 15
 ADR00645
 ID ADR00645 standard; peptide; 9 AA.

XX AC ADR00645;

XX DT 04-NOV-2004 (first entry)

XX DE Human 254P1D6B peptide SEQ ID NO:56.

XX KW 254P1D6B; small interfering RNA; siRNA; immune response;
 XX KW 254P1D6B-related protein; cytostatic; gene therapy; cancer; human;
 XX KW chromosome 6.

XX OS Homo sapiens.

XX PN WO2004067716-A2.

XX PD 12-AUG-2004.

XX PF 23-JAN-2004; 2004WO-US001965.

XX PR 24-JAN-2003; 2003US-0442526P.

XX PA (AGEN-) AGENSYS INC.

XX PI Kanner SB, Raitano AB, Jakobovits A, Challita-Eid PM, Ge W;
 XX PI Perez-Villar JJ, Farris M;

XX DR WPI; 2004-580991/56.

XX PT New 254P1D6B siRNA composition comprising a double stranded siRNA that
 XX PT corresponds to the nucleic acid ORF sequence which encodes the 254P1D6B
 XX PT protein or corresponds to a subsequence of the ORF, useful for detecting
 XX PT and treating cancer.

XX PS Example 9; SEQ ID NO 56; 345pp; English.

XX CC The present invention describes a 254P1D6B small interfering RNA (siRNA)
 CC composition that comprises a double stranded siRNA that corresponds to
 CC the nucleic acid open reading frame (ORF) sequence which encodes the
 CC 254P1D6B protein, or corresponds to a subsequence of the ORF, where the
 CC double stranded siRNA is 19, 20, 21, 22, 23, 24, or 25 contiguous
 CC nucleotides in length. Also described: (1) a composition that comprises,
 CC consists essentially of, or consists of a peptide of eight, nine, ten, or

CC eleven contiguous amino acids of a protein of figure 2 (P1, see SEQ ID
 CC NO:13, 5 or 7 ADR00592, ADR00594 or ADR00596), or a peptide included in
 CC any of the 42 lists of peptides, given in the specification, or a protein
 CC that is at least 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% homologous or
 CC identical to an entire amino acid sequence of P1; (2) a polynucleotide
 CC that encodes the protein; (3) a composition comprising a polynucleotide
 CC that is a full complement of the polynucleotide described above; (4)
 CC generating a mammalian immune response directed to the protein of P1; (5)
 CC detecting, in a sample, the presence of a 254P1D6B-related protein or a
 CC 254P1D6B-related polynucleotide; (6) a composition that modulates the
 CC status of a cell that expresses a protein of P1; (7) a pharmaceutical
 CC composition that comprises the composition described above in a human
 CC unit dose form; (8) an antibody or its fragment, which is monoclonal; (9)
 CC a non-human transgenic animal that produces the antibody; (10) a
 CC hybridoma that produces the antibody; (11) delivering a cytotoxic agent
 CC or a diagnostic agent to a cell that expresses the protein of P1; and
 CC (12) inhibiting growth, reproduction or survival of cancer cells that
 CC express the protein of P1. 254P1D6B has cytostatic activity, and can be
 CC used in gene therapy. The compositions, molecules and methods are useful
 CC for treating and detecting cancer. The present sequence represents the
 CC human 254P1D6B peptide, which is used in the exemplification of the
 CC present invention. The human 254P1D6B gene is located on chromosome 6p22.
 XX

Sequence 9 AA;

Query Match 0.8%; Score 9; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 IWGDSPEDI 137
 Db 1 IWGDSPEDI 9
 |||||

Search completed: October 12, 2005, 10:24:01
 Job time : 74 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:27:56 ; Search time 70 Seconds

(without alignments)
6370.973 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 5580

Sequence: 1 MAPRTGVLSLLLVITAGC.....VSMNGSIRNGASRYSCKDR 1072

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1854112 seqs, 416015017 residues

Total number of hits satisfying chosen parameters: 458990

Minimum DB seq length: 0

Maximum DB seq length: 35

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158	2.8	29	16 US-10-764-390-268	Sequence 268, App
2	152	2.7	29	16 US-10-764-390-262	Sequence 262, App
3	101	1.8	19	16 US-10-764-390-261	Sequence 261, App
4	99	1.8	19	16 US-10-764-390-267	Sequence 267, App
5	92	1.6	16	16 US-10-764-390-43	Sequence 43, App
6	90	1.6	17	16 US-10-764-390-266	Sequence 266, App
7	84	1.5	17	16 US-10-764-390-260	Sequence 260, App
8	83	1.5	19	18 US-10-450-763-52425	Sequence 52425, A
9	79	1.4	24	16 US-10-764-390-265	Sequence 265, App
10	63	1.1	9	16 US-10-764-390-138	Sequence 138, App
11	61.5	1.1	34	9 US-09-892-835-1	Sequence 1, Appli

12	61.5	1.1	34	10	US-09-769-180-3	Sequence 3, Appli
13	61.5	1.1	34	15	US-10-682-103-3	Sequence 3, Appli
14	60	1.1	12	16	US-10-764-390-41	Sequence 41, Appl
15	59	1.1	9	16	US-10-764-390-137	Sequence 137, App
16	59	1.1	9	16	US-10-764-390-249	Sequence 249, App
17	57	1.0	9	16	US-10-764-390-52	Sequence 52, Appl
18	57	1.0	31	10	US-09-984-130-64	Sequence 64, Appl
19	57	1.0	31	10	US-09-836-353A-64	Sequence 64, Appl
20	57	1.0	34	14	US-10-029-386-30913	Sequence 30913, A
21	56	1.0	9	16	US-10-764-390-139	Sequence 139, App
22	56	1.0	9	16	US-10-764-390-225	Sequence 225, App
23	56	1.0	9	16	US-10-764-390-240	Sequence 240, App
24	55	1.0	9	16	US-10-764-390-50	Sequence 50, Appl
25	55	1.0	9	16	US-10-764-390-86	Sequence 86, Appl
26	55	1.0	9	16	US-10-764-390-239	Sequence 239, App
27	55	1.0	11	16	US-10-764-390-42	Sequence 42, Appl
28	55	1.0	27	15	US-10-038-854-390	Sequence 390, App
29	55	1.0	32	14	US-10-062-831-177	Sequence 177, App
30	55	1.0	32	14	US-10-062-599-177	Sequence 177, App
31	55	1.0	33	10	US-09-984-130-58	Sequence 58, Appl
32	55	1.0	33	10	US-09-836-353A-58	Sequence 58, Appl
33	54	1.0	9	16	US-10-764-390-105	Sequence 105, App
34	54	1.0	9	16	US-10-764-390-136	Sequence 136, App
35	54	1.0	9	16	US-10-764-390-158	Sequence 158, App
36	54	1.0	9	16	US-10-764-390-221	Sequence 221, App
37	54	1.0	9	16	US-10-764-390-237	Sequence 237, App
38	54	1.0	29	15	US-10-176-419A-4	Sequence 4, Appli
39	54	1.0	34	10	US-09-984-130-62	Sequence 62, Appl
40	54	1.0	34	10	US-09-836-353A-62	Sequence 62, Appl
41	53.5	1.0	24	14	US-10-292-896-12	Sequence 12, Appl
42	53	0.9	9	16	US-10-764-390-55	Sequence 55, Appl
43	53	0.9	9	16	US-10-764-390-56	Sequence 56, Appl
44	53	0.9	9	16	US-10-764-390-95	Sequence 95, Appl
45	53	0.9	9	16	US-10-764-390-177	Sequence 177, App

ALIGNMENTS

RESULT 1

US-10-764-390-268
; Sequence 268, Application US/10764390
; Publication No. US2004021421A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Rid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 268
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-268

Query Match 2.8%; Score 158; DB 16; Length 29;
Best Local Similarity 96.6%; Pred. No. 0.0062;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 128 GINGSPEDIRKDLXFLGKWGLEMSEY 156

```
Db 1 GIGDSPDIRKDLTFLGKDWGLEMSEY 29
|||||
RESULT 2
US-10-764-390-262
; Sequence 262, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 262
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-262

Query Match 2.7%; Score 152; DB 16; Length 29;
Best Local Similarity 96.6%; Pred. No. 0.015;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 143 FLGKDWGLEMSEYDDYRELEKDLQPS 171
|||||
Db 1 FLGKDWGLEMSEYADDYRELEKDLQPS 29
|||||

RESULT 3
US-10-764-390-261
; Sequence 261, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 261
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-261

Query Match 1.8%; Score 101; DB 16; Length 19;
Best Local Similarity 94.7%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 148 WGLEMSEYDDYRELEKD 166
|||||
Db 1 WGLEMSEYADDYRELEKD 19
|||||

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 148 WGLEMSEYDDYRELEKD 166
|||||
Db 1 WGLEMSEYADDYRELEKD 19
|||||

RESULT 4
US-10-764-390-267
; Sequence 267, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 267
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-267

Query Match 1.8%; Score 99; DB 16; Length 19;
Best Local Similarity 94.7%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 133 SPEDIRKDLTFLGKDWGLE 151
|||||
Db 1 SPEDIRKDLTFLGKDWGLE 19
|||||

RESULT 5
US-10-764-390-43
; Sequence 43, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-43
```

Query Match 1.6%; Score 92; DB 16; Length 16;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 522 NNAVDPYPVANAGPNH 537
Db 1 NNAVDPYPVANAGPNH 16

RESULT 6

US-10-764-390-266
; Sequence 266, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 266
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-266

Query Match 1.6%; Score 90; DB 16; Length 17;
Best Local Similarity 94.1%; Pred. No. 45;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 134 PEDIRKDLXFLGKDWGL 150
Db 1 PEDIRKDLXFLGKDWGL 17

RESULT 7

US-10-764-390-260
; Sequence 260, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260
; LENGTH: 17

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-260

Query Match 1.5%; Score 84; DB 16; Length 17;
Best Local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 149 GLEEMSEYDDYRELEK 165
Db 1 GLEEMSEYDDYRELEK 17

RESULT 8

US-10-450-763-52425
; Sequence 52425, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 03/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 03/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 52425
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(19)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-52425

Query Match 1.5%; Score 83; DB 18; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPPTGVLSLLLVTTIA 18
Db 1 MAPPTGVLSLLLVTTIA 18

RESULT 9

US-10-764-390-265
; Sequence 265, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0


```
; SEQ ID NO 265
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-265

Query Match      1.4%; Score 79; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 CARQCSSEGRVSN 33
    |||||
Db 11 CARQCSSEGRVSN 24
    |||||

RESULT 10
US-10-764-390-138
; Sequence 138, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Paris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254p1b6B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 138
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-138

Query Match      1.1%; Score 63; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 946 WDGESNCEW 954
    |||||
Db 1 WDGESNCEW 9

RESULT 11
US-09-892-835-1
; Sequence 1, Application US/09892835
; Publication No. US2002019016A1
; GENERAL INFORMATION:
; APPLICANT: Hulstaert, Frank
; APPLICANT: Vanmechelen, Eugene
; APPLICANT: Vanderstichele, Hugo
; TITLE OF INVENTION: Differential Diagnosis of Neurological Diseases
; FILE REFERENCE: INNS:027 (11362.0027.NPUS01)
; CURRENT APPLICATION NUMBER: US/09/892,835
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 09/892.835
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/218.907
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: EP 00870151.8
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; NAME/KEY: MOD RES
; LOCATION: (28)..(28)
; OTHER INFORMATION: PHOSPHORYLATION THREONINE
US-09-892-835-1

Query Match      1.1%; Score 61.5; DB 9; Length 34;
Best Local Similarity 29.6%; Pred. No. 6.8e+03;
Matches 21; Conservative 6; Mismatches 7; Indels 37; Gaps 4;

QY 176 PRGSAEYTDWGLLPQSEGAFFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVL 235
    |||||
Db 1 PRGAAP-----PQKQGANAT-----RIPAKT-----PPAPKTP----- 29
    |||||

QY 236 LPLPTTPSSGE 246
    |||||
Db 30 -----PSSGE 34

RESULT 12
US-09-769-180-3
; Sequence 3, Application US/09769180
; Publication No. US20030194742A1
; GENERAL INFORMATION:
; APPLICANT: VanMechelen, Eugene
; APPLICANT: Vanderstichele, Hugo
; TITLE OF INVENTION: Diagnosis of Tauopathies
; FILE REFERENCE: US.112.T181
; CURRENT APPLICATION NUMBER: US/09/769,180
; CURRENT FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: EP 00870008.8
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: EP 00870280.5
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/178,391
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (28)..(28)
; OTHER INFORMATION: PHOSPHORYLATION
US-09-769-180-3

Query Match      1.1%; Score 61.5; DB 10; Length 34;
Best Local Similarity 29.6%; Pred. No. 6.8e+03;
Matches 21; Conservative 6; Mismatches 7; Indels 37; Gaps 4;

QY 176 PRGSAEYTDWGLLPQSEGAFFNSVGDSPAVPAETQDDPELHYLNESASTPAPKLPERSVL 235
    |||||
Db 1 PRGAAP-----PQKQGANAT-----RIPAKT-----PPAPKTP----- 29
    |||||

QY 236 LPLPTTPSSGE 246
    |||||
Db 30 -----PSSGE 34

RESULT 13
US-10-682-103-3
; Sequence 3, Application US/10682103
; Publication No. US20040091942A1
; GENERAL INFORMATION:
; APPLICANT: VanMechelen, Eugene
```

```
; APPLICANT: Vanderstichele, Hugo
; TITLE OF INVENTION: Diagnosis of Taupathies
; FILE REFERENCE: 11916.0016.DVUS01 (INNS016--1)
; CURRENT APPLICATION NUMBER: US/10/682,103
; CURRENT FILING DATE: 2003-10-09
; PRIOR FILING DATE: 2003-10-09
; PRIOR FILING DATE: 09/769,180
; PRIOR APPLICATION NUMBER: EP 00870008.8
; PRIOR FILING DATE: 2000-01-24
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: EP 00870280.5
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/178,391
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (28)..(28)
; OTHER INFORMATION: PHOSPHORYLATION
US-10-682-103-3

Query Match      1.1%; Score 61.5; DB 15; Length 34;
Best Local Similarity 29.6%; Pred. No. 6.8e+03;
Matches 21; Conservative 6; Mismatches 7; Indels 37; Gaps 4;

Qy 176 PRGSAEYTDGLLPGSGAFNSVGDSPAVPAETQQDPQLHYLNESASTAPKLPERSVL 235
Db 1 PRGAAP-----PQKGQANAT-----RIPAKT-----PPAKPTP----- 29

Qy 236 LPLPTPSSGE 246
Db 30 -----PSSGE 34

RESULT 14
US-10-764-390-41
; Sequence 41, Application US/10764390
; Publication No. US20040214212A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-41

Query Match      1.1%; Score 60; DB 16; Length 12;
Best Local Similarity 91.7%; Pred. No. 2e+03;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 151 REMSEYDDYRE 162
Db 151 REMSEYDDYRE 162

; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Ge, Wangmao
; APPLICANT: Faris, Mary
; APPLICANT: Steven B. Kanner
; APPLICANT: Juan J. Perez-Villar
; TITLE OF INVENTION: Nucleic Acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 254PID6B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20081.00
; CURRENT APPLICATION NUMBER: US/10/764,390
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US60/442,526
; PRIOR FILING DATE: 2003-01-24
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 137
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-764-390-137

Query Match      1.1%; Score 59; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 929 RCICSHLWM 937
Db 1 RCICSHLWM 9

Search completed: October 12, 2005, 10:33:06
Job time : 70 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:26:35 ; Search time 26 Seconds
(without alignments)
3077.838 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 5580
Sequence: 1 MAPPTGVLSLLLVITAGC.....VSMNGSIRNGASFSYCSKDR 1072

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 250370

Minimum DB seq length: 0
Maximum DB seq length: 35

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61.5	1.1	34	4	US-09-892-835-1
2	61.5	1.1	34	4	US-09-769-180-3
3	59.5	1.1	35	3	US-08-658-136-52
4	55	1.0	32	4	US-09-690-454-177
5	54	1.0	25	4	US-09-060-767B-4
6	53.5	1.0	24	4	US-09-217-306B-19
7	51	0.9	31	3	US-08-602-999A-54
8	51	0.9	31	3	US-08-278-865-54
9	51	0.9	31	4	US-09-500-124-54
10	51	0.9	31	4	US-09-938-315-54
11	49	0.9	22	3	US-09-430-564-9
12	49	0.9	22	3	US-09-430-564-15
13	49	0.9	23	2	US-08-833-807-6
14	49	0.9	23	3	US-09-233-043-6
15	49	0.9	23	4	US-09-593-870A-6
16	49	0.9	30	3	US-09-039-780A-87
17	49	0.9	31	2	US-08-244-951A-4
18	49	0.9	31	3	US-09-280-047-10
19	49	0.9	31	3	US-08-208-573B-10
20	49	0.9	33	2	US-08-389-011-2
21	49	0.9	33	3	US-08-403-917A-2
22	49	0.9	33	3	US-09-348-952A-2
23	49	0.9	34	2	US-08-602-264A-10
24	49	0.9	34	3	US-08-461-018A-10
25	49	0.9	34	3	US-09-216-958-10
26	49	0.9	35	2	US-08-244-951A-6
27	48.5	0.9	18	4	US-09-910-009A-401

28 48 0.9 17 4 US-09-910-009A-395 Sequence 395, App
29 48 0.9 17 4 US-09-910-009A-413 Sequence 413, App
30 48 0.9 17 4 US-09-910-009A-447 Sequence 447, App
31 48 0.9 18 4 US-09-910-009A-451 Sequence 451, App
32 48 0.9 21 4 US-09-646-154-1 Sequence 1, Appli
33 48 0.9 27 2 US-08-394-600B-4 Sequence 4, Appli
34 48 0.9 27 4 US-08-385-456C-4 Sequence 4, Appli
35 48 0.9 27 4 US-08-487-453A-4 Sequence 4, Appli
36 48 0.9 27 5 PCT-US95-02513-4 Sequence 9, Appli
37 48 0.9 30 3 US-09-050-811-9 Sequence 16, Appl
38 48 0.9 34 2 US-08-942-423-16 Sequence 16, Appl
39 47 0.8 27 4 US-09-636-399A-17 Sequence 17, Appl
40 47 0.8 29 4 US-09-636-399A-15 Sequence 15, Appl
41 47 0.8 30 4 US-09-636-399A-16 Sequence 16, Appl
42 47 0.8 30 4 US-09-636-399A-72 Sequence 72, Appl
43 47 0.8 30 4 US-09-270-767-59044 Sequence 59044, A
44 47 0.8 30 4 US-09-270-767-52474 Sequence 62474, A
45 47 0.8 31 4 US-09-636-399A-70 Sequence 70, Appl

ALIGNMENTS

RESULT 1
US-09-892-835-1
; Sequence 1, Application US/09892835
; Patent No. 6670137
; GENERAL INFORMATION:
; APPLICANT: Hulstaert, Frank
; APPLICANT: Vanmechelen, Eugene
; APPLICANT: Vanderstichele, Hugo
; TITLE OF INVENTION: Differential diagnosis of Neurological Diseases
; FILE REFERENCE: INNS:027 (11362.0027.NPUS01)
; CURRENT APPLICATION NUMBER: US/09/892.835
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 09/892.835
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/218.907
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: EP 00870151.8
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (28)..(28)
; OTHER INFORMATION: PHOSPHORYLATION THREONINE
US-09-892-835-1

Query Match 1.1%; Score 61.5; DB 4; Length 34;
Best Local Similarity 29.6%; Pred. No. 4.1e+02;
Matches 21; Conservative 6; Mismatches 7; Indels 37; Gaps 4;

Qy 176 PRGSAEYTDGLPLPGSEGFNSVGDSPVAPAEATQDDPDLHYLINESASTPAPKLPERSVL 235
Db 1 PRGAAP-----PGQKQGANAT-----RIPAKT-----PPAPKTFP----- 29
Qy 236 LPLPTTPSSGE 246
Db 30 -----PSSGE 34

RESULT 2
US-09-769-180-3
; Sequence 3, Application US/09769180
; Patent No. 6680173
; GENERAL INFORMATION:

```

; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-658-136-52

Query Match 1.1%; Score 59.5; DB 3; Length 35;
Best Local Similarity 35.0%; Pred. No. 6.2e+02;
Matches 14; Conservative 3; Mismatches 18; Indels 5; Gaps 1;

QY 560 VLYEWSLGPSSGKHVVMQGVQTPYLHLSAMQEGDYTFQL 599
      |||||
Db 1 VAYHWDFGDGSFG-----QDTDEPRAEHSYLRGEGYRVQV 35
      |||||

RESULT 4
US-09-690-454-177
; Sequence 177, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 177
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-690-454-177

Query Match 1.0%; Score 55; DB 4; Length 32;
Best Local Similarity 45.2%; Pred. No. 1.2e+03;
Matches 14; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 155 EYXDDYRELEKOLLQSPGKQEPGRGSABYTDW 185
      |||||
Db 1 EYNLDYTELGLQKLKESGKQ--HGFASFSDY 29
      |||||

RESULT 5
US-09-060-767B-4
; Sequence 4, Application US/09060767B
; Patent No. 6720152
; GENERAL INFORMATION:
; APPLICANT: Weill, Gary
; APPLICANT: Chandrashekar, Ramaswamy
; TITLE OF INVENTION: Diagnosis of Histoplasmosis Using Antigens Specific for

```

;; TITLE OF INVENTION: H. capsulatum
;; FILE REFERENCE: BJCH 9986
;; CURRENT APPLICATION NUMBER: US/09/060.767B
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/043,332
;; PRIOR FILING DATE: 1997-04-15
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 4
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Caldocellum saccharolyticum
US-09-060-767B-4

Query Match 1.0%; Score 54; DB 4; Length 25;
Best Local Similarity 45.8%; Pred. No. 9.8e+02;
Matches 11; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

Qy 291 PVLTVPGSTHSIPTPTSAAPS 314
Db 1 PTSTVTPTPTPTPTPTPTPTPT 24

RESULT 6
US-09-217-306B-19
; Sequence 19, Application US/09217306B
; Patent No. 6465220
; GENERAL INFORMATION:
; APPLICANT: Hassan, Helle
; APPLICANT: Clausen, Henrik
; APPLICANT: Bennett, Eric P.
; TITLE OF INVENTION: Glycosylation Using GalNAc-T4 Transferase
; FILE REFERENCE: 8850*1
; CURRENT APPLICATION NUMBER: US/09/217.306B
; CURRENT FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; OTHER INFORMATION: Muc7
US-09-217-306B-19

Query Match 1.0%; Score 53.5; DB 4; Length 24;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

Qy 296 TPGSTHSIPTPTSAAPSEST 317
Db 4 TPSAT---TPAPPSSAPPETT 22

RESULT 7
US-08-602-999A-54
; Sequence 54, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/602.999A
;; FILING DATE: 16-FEB-1996
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Misrock, S. Leslie
;; REGISTRATION NUMBER: 18,872
;; REFERENCE/DOCKET NUMBER: 1101-202
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-9741/8864
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 54:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 31 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-602-999A-54

Query Match 0.9%; Score 51; DB 3; Length 31;
Best Local Similarity 62.5%; Pred. No. 2.4e+03;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 314 SESTPSELPISTPTAP 329
Db 14 SRSTPRLPMLPTTRP 29

RESULT 8
US-08-278-865-54
; Sequence 54, Application US/08278865
; Patent No. 6303574
; GENERAL INFORMATION:
; APPLICANT: KAY, BRIAN K.
; APPLICANT: SPARKS, ANDREW B.
; APPLICANT: THORN, JUDITH M.
; APPLICANT: QUILLIAM, LAWRENCE A.
; APPLICANT: DER, CHANNING J.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/278,865
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Villacorta, Gilberto M.
;; REGISTRATION NUMBER: 34,038
;; REFERENCE/DOCKET NUMBER: 4980-007-0
;; TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-278-865-54

Query Match 0.9%; Score 51; DB 3; Length 31;
Best Local Similarity 62.5%; Pred. No. 2.4e+03;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 314 SESTPSELPISTTAP 329
DB 14 SRSTPRPLMLPTTRP 29

RESULT 9
US-09-500-124-54
Sequence 54, Application US/09500124
Patent No. 6432920
GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: QUILLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLKES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
TITLE OF INVENTION: ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/500,124
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/602,999
FILING DATE: 16-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-500-124-54

Query Match 0.9%; Score 51; DB 4; Length 31;
Best Local Similarity 62.5%; Pred. No. 2.4e+03;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 314 SESTPSELPISTTAP 329
DB 14 SRSTPRPLMLPTTRP 29

RESULT 10
US-09-938-315-54
Sequence 54, Application US/09938315
Patent No. 6703482
GENERAL INFORMATION:

APPLICANT: KAY, BRIAN K.
SPARKS, ANDREW B.
THORN, JUDITH M.
QUILLIAM, LAWRENCE A.
DER, CHANNING J.

TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
ISOLATING AND USING SAME

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/938,315

FILING DATE: 23-Aug-2001

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Villacorta, Gilberto M.

REGISTRATION NUMBER: 34,038

REFERENCE/DOCKET NUMBER: 4980-007-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 54:

US-09-938-315-54

Query Match 0.9%; Score 51; DB 4; Length 31;
Best Local Similarity 62.5%; Pred. No. 2.4e+03;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 314 SESTPSELPISTTAP 329
DB 14 SRSTPRPLMLPTTRP 29

RESULT 11
US-09-430-564-9
Sequence 9, Application US/09430564
Patent No. 6372467
GENERAL INFORMATION:

APPLICANT: John Blenis

APPLICANT: Kay K. Lee-Fruman

APPLICANT: Calvin J. Kuo

TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,

TITLE OF INVENTION: PRIMERS, PROBES, AND DETECTION METHODS

FILE REFERENCE: 00246/506002

```
; CURRENT APPLICATION NUMBER: US/09/430,564
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,141
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-430-564-9

Query Match      0.9%  Score 49;  DB 3;  Length 22;
Best Local Similarity 47.6%;  Pred. No. 2e+03;
Matches 10;  Conservative 1;  Mismatches 10;  Indels 0;  Gaps 0;

Qy  305  PTPPTSAAPSESTPSELPISP 325
Db  1  PLPPLPPPPPTTAPLPFIRP 21

RESULT 12
US-09-430-564-15
; Sequence 15, Application US/09430564
; Patent No. 6372467
; GENERAL INFORMATION:
; APPLICANT: John Blenis
; APPLICANT: Kay K. Lee-Fruman
; APPLICANT: Calvin J. Kuo
; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
; TITLE OF INVENTION: PRIMERS, PROBES, AND DETECTION METHODS
; FILE REFERENCE: 00246/506002
; CURRENT APPLICATION NUMBER: US/09/430,564
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,141
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-430-564-15

Query Match      0.9%  Score 49;  DB 3;  Length 22;
Best Local Similarity 47.6%;  Pred. No. 2e+03;
Matches 10;  Conservative 1;  Mismatches 10;  Indels 0;  Gaps 0;

Qy  305  PTPPTSAAPSESTPSELPISP 325
Db  1  PLPPLPPPPPTTAPLPFIRP 21

RESULT 13
US-08-833-807-6
; Sequence 6, Application US/08833807
; Patent No. 5989552
; GENERAL INFORMATION:
; APPLICANT: McKenzie, Ian F.C.
; APPLICANT: Apostolopoulos, Vasso
; APPLICANT: Pietersz, Geoff A.
; TITLE OF INVENTION: ANTIGENIC CARBOHYDRATE COMPOUNDS AND
; TITLE OF INVENTION: THEIR USE IN IMMUNOTHERAPY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann Dorfman Herrell and Skillman
; STREET: Suite 720, 1601 Market Street
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: United States of America
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/223,043
; APPLICATION NUMBER: US/09/223,043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/833,807
; FILING DATE:
; APPLICATION NUMBER: AU PM3223
; FILING DATE: 24-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagen, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 530547/PAS/MKR

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,807
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/833,807
; FILING DATE:
; APPLICATION NUMBER: AU PM3223
; FILING DATE: 24-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagen, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 530547/PAS/MKR

; Query Match      0.9%  Score 49;  DB 2;  Length 23;
; Best Local Similarity 40.9%;  Pred. No. 2.2e+03;
; Matches 9;  Conservative 4;  Mismatches 9;  Indels 0;  Gaps 0;

Qy  294  TVTPGSTSHSIPTPTPSAAPSE 315
Db  2  TAAPTPTTATPPPPSSAPPE 23

RESULT 14
US-09-223-043-6
; Sequence 6, Application US/09223043
; Patent No. 6177256
; GENERAL INFORMATION:
; APPLICANT: McKenzie, Ian F.C.
; APPLICANT: Apostolopoulos, Vasso
; APPLICANT: Pietersz, Geoff A.
; TITLE OF INVENTION: ANTIGENIC CARBOHYDRATE COMPOUNDS AND
; TITLE OF INVENTION: THEIR USE IN IMMUNOTHERAPY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann Dorfman Herrell and Skillman
; STREET: Suite 720, 1601 Market Street
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: United States of America
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/223,043
; APPLICATION NUMBER: US/09/223,043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/833,807
; FILING DATE:
; APPLICATION NUMBER: AU PM3223
; FILING DATE: 24-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagen, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 530547/PAS/MKR
```


TELECOMMUNICATION INFORMATION:
TELEPHONE: (215)563-4100
TELEFAX: (215)563-4044
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-223-043-6

Query Match 0.9%; Score 49; DB 3; Length 23;
Best Local Similarity 40.9%; Pred. No. 2.2e+03;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 294 TVTPGSTSHSIPTPTSAAPSE 315
Db 2 TAAPPTTPATTAPPSAPPE 23

RESULT 15
US-09-593-870A-6
Sequence 6, Application US/09593870A
Patent No. 6548643
GENERAL INFORMATION:
APPLICANT: McKenzie, Ian F. C.
APPLICANT: Apostolopoulos, Vasso
APPLICANT: Pietersz, Geoff Allan
TITLE OF INVENTION: Antigen Carbohydrate Compounds and Their
FILE REFERENCE: 2368-McKenzie
CURRENT APPLICATION NUMBER: US/09/593,870A
CURRENT FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: 09/223,043
PRIOR FILING DATE: 1998-12-30
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 23
TYPE: PRT
ORGANISM: Homo sapiens
US-09-593-870A-6

Query Match 0.9%; Score 49; DB 4; Length 23;
Best Local Similarity 40.9%; Pred. No. 2.2e+03;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 294 TVTPGSTSHSIPTPTSAAPSE 315
Db 2 TAAPPTTPATTAPPSAPPE 23

Search completed: October 12, 2005, 10:31:46
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:25:25 ; Search time 67 Seconds

(without alignments)
8193.267 Million cell updates/sec

Title: US-10-764-390-3

Perfect score: 5580

Sequence: 1 MAPPTGVLSLLLVTTIAGC.....VSMNGSIRNGASPSYCSKDR 1072

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 37082

Minimum DB seq length: 0

Maximum DB seq length: 35

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	51	0.9	32	Q7M4E6	Q7M4E6 drosophila
2	49.5	0.9	34	Q6PJK3	Q6PJK3 homo sapien
3	48	0.9	27	Q8CLU0	Q8CLU0 yersinia pe
4	48	0.9	35	1 TXH5 ORNHU	P61104 ornithocton
5	47.5	0.9	27	1 TXA3 ANESU	P01535 anemonia su
6	47	0.8	18	2 QSLD14	Q61D14 homo sapien
7	47	0.8	21	2 Q8WLR1	Q8WLR1 homo sapien
8	47	0.8	23	1 SKL1_TITCA	P83243 titiys camb
9	46	0.8	24	2 Q7YP58	Q7YP58 homo sapien
10	46	0.8	29	2 Q6WY52	Q6WY52 bacterioph
11	46	0.8	32	1 TAT_SIVM2	P05912 simian immu
12	46	0.8	34	2 Q70222	Q70222 human immu
13	45.5	0.8	35	2 Q9YMB1	Q9YMB1 human immu
14	45	0.8	27	2 Q9Y181	Q9Y181 priapulid c
15	45	0.8	29	2 Q7R7E1	Q7R7E1 plasmidom
16	45	0.8	33	2 Q87JB9	Q87JB9 vibrio para
17	45	0.8	33	2 Q39816	Q39816 encephalomy
18	44.5	0.8	34	2 Q735G0	Q735G0 bacillus ce
19	44.5	0.8	35	2 Q80586	Q80586 human immu
20	44	0.8	34	2 Q8J3X2	Q8J3X2 human immu
21	44	0.8	35	1 SCXP_ANDMA	P01498 androctonus
22	43.5	0.8	20	2 Q8IXG1	Q8IXG1 homo sapien
23	43.5	0.8	33	2 Q8EFK7	Q8EFK7 shewanella
24	43.5	0.8	35	2 Q74EN3	Q74EN3 geobacter s
25	43	0.8	22	2 Q85607	Q85607 streptomyce
26	43	0.8	26	2 Q38675	Q38675 bacterioph
27	43	0.8	27	2 Q77D54	Q77D54 bacterioph
28	43	0.8	27	2 Q9AZ16	Q9AZ16 bacterioph
29	43	0.8	34	2 Q7Y043	Q7Y043 vitis vinif
30	43	0.8	34	2 Q2ZG81	Q2ZG81 chlamydia t
31	43	0.8	35	2 Q83333	Q83333 murine hepa

RESULT 1

ID	Q7M4E6	PRELIMINARY;	PRT;	32 AA.
AC	Q7M4E6;			
DT	01-MAR-2004	(TrEMBLrel. 26, Created)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DE	Translation initiation factor eIF-4E (Fragments).			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE.			
RA	Hernandez G., Sierra J.M.;			
RT	"Translation initiation factor eIF-4E from Drosophila: cDNA sequence			
RT	and expression of the gene.";			
RL	Biochim. Biophys. Acta 1261:427-431 (1995).			
DR	PIR, S55937; S55937.			
FT	NON TER 32			
SQ	SEQUENCE 32 AA; 3364 MW; C492C82B551A6959 CRC64;			

Query Match 0.9%; Score 51; DB 2; Length 32;
Best Local Similarity 69.8%; Pred. No. 3e+04;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy	299	STHSIPTPTPTSAAPS	314
Db	13	STEQGRPEPTSAAS	28

RESULT 2

ID	Q6PJK3	PRELIMINARY;	PRT;	34 AA.
AC	Q6PJK3;			
DT	05-JUL-2004	(TrEMBLrel. 27, Created)		
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)		
DE	PMW1G protein.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Scheetz T.E.,			

32	42.5	0.8	32	1	RK1_RABIT	P81655 oryctolagus
33	42.5	0.8	34	2	Q7RK45	Q7RK45 plasmodium
34	42.5	0.8	35	2	Q7RX56	Q7RX56 neurospora
35	42.5	0.8	35	2	Q9Y083	Q9Y083 globodera r
36	42.5	0.8	35	2	Q77900	Q77900 human immu
37	42.5	0.8	35	2	Q80555	Q80555 human immu
38	42.5	0.8	35	2	Q80562	Q80562 human immu
39	42.5	0.8	35	2	Q80563	Q80563 human immu
40	42.5	0.8	35	2	Q80564	Q80564 human immu
41	42.5	0.8	35	2	Q80565	Q80565 human immu
42	42	0.8	22	2	Q7M3F1	Q7M3F1 bos taurus
43	42	0.8	23	2	Q8HA12	Q8HA12 bacterioph
44	42	0.8	24	1	LPER_STRFR	P45440 streptomyce
45	42	0.8	25	2	Q7Y286	Q7Y286 phage phi 4

ALIGNMENTS

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014386; AAH14386.2; -
SQ SEQUENCE 34 AA; 3886 MW; 7B38FEFC8B7809D4 CRC64;

Query Match 0.9%; Score 48.5; DB 2; Length 34;
Best Local Similarity 37.0%; Pred. No. 4.4e+04;
Matches 10; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

QY 915 LKCSGHGCHDPLTKRCICSHLWMENLI 941
Db 1 MKMGSGFGYCHPLWKSCWIS-AWHQTLIL 26

RESULT 3
Q8CLU0 PRELIMINARY; PRT; 27 AA.
ID Q8CLU0
AC Q8CLU0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical.
GN OrderedLocusNames=y0241;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.V.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AF013624; AAM83835.1; -.
KW Hypothetical protein.
SQ SEQUENCE 27 AA; 3211 MW; A5203621B3DE0ABC CRC64;

Query Match 0.9%; Score 48; DB 2; Length 27;
Best Local Similarity 39.1%; Pred. No. 3.4e+04;
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 841 LAVLLNVLPSDIKVKQIRAHSDL 863
Db 4 LLLLIQRTDSEIKLVRGSHSL 26

RESULT 4
TXH5_ORNHU
ID TXH5_ORNHU
AC P61104; PRT; 35 AA.

```

```

DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Huwentoxin-V (HwTx-V) [Contains: Mutant of huwentoxin-V (mHwTx-V)].
OS Ornithotoxus huwena (Chinese bird spider) (Selenocosmia huwena).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Mygalomorphae; Theraphosidae; Ornithotoxus.
OX NCBI_TaxID=29017;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX PubMed=12893056; DOI=10.1016/S0041-0101(03)00095-3;
RA Zhang P.-F., Chen P., Hu W.-J., Liang S.-P.;
RT "Huwentoxin-V, a novel insecticidal peptide toxin from the spider
RT Selenocosmia huwena, and a natural mutant of the toxin: indicates the
RT key amino acid residues related to the biological activity.";
RL Toxicon 42:15-20(2003).
CC -!- FUNCTION: HwTx-V reversibly paralyzes locusts and cockroaches, and
CC causes death with larger doses. It has no effect on mice by intra-
CC abdominal, nor intracerebroventricular injection. Its natural
CC mutant mHwTx-V shows no effect on locusts, cockroaches, and mice.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- MASS SPECTROMETRY: MW=4111.4; MW_ERR=0.4; METHOD=MALDI; RANGE=1-
CC 35; NOTE=Ref.1.
CC -!- MASS SPECTROMETRY: MW=3877.1; MW_ERR=0.4; METHOD=MALDI; RANGE=1-
CC 33; NOTE=Ref.1.
CC -!- MISCELLANEOUS: ED(50) of HwTx-V is 16 +/- 5 mg/kg to locusts.
KW Direct protein sequencing; Neurotoxin; Toxin.
FT CHAIN 1 35 Huwentoxin-V.
FT CHAIN 1 33 Mutant of huwentoxin-V.
FT DISULFID 2 16 By similarity.
FT DISULFID 9 21 By similarity.
FT DISULFID 15 28 By similarity.
SQ SEQUENCE 35 AA; 4117 MW; CC39F421571CA523 CRC64;

Query Match 0.9%; Score 48; DB 1; Length 35;
Best Local Similarity 31.2%; Pred. No. 4.8e+04;
Matches 10; Conservative 3; Mismatches 11; Indels 8; Gaps 1;

QY 917 CSGHGHCDPLTKRCICSHLWMENLIQRYIWDG 948
Db 9 CSQGDGDC-----CKHLQCHSNYWCWVDG 32

RESULT 5
TXA3_ANESU
ID TXA3_ANESU STANDARD; PRT; 27 AA.
AC P01535;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Neurotoxin III (Toxin ATX-III).
OS Anemonia sulcata (Snake-locks sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthaeae; Actiniidae; Anemonia.
OX NCBI_TaxID=6108;
RN [1]
RP SEQUENCE.
RC TISSUE=Cnidoblast;
RX MEDLINE=78084776; PubMed=23311; DOI=10.1016/0014-5793(77)80699-6;
RA Martinez G., Kopeyan C., Schweitz H., Lazdunski M.;
RT "Toxin III from Anemonia sulcata: primary structure.";
RL FEBS Lett. 84:247-252(1977).
RN [2]
RP PRELIMINARY SEQUENCE.
RX MEDLINE=78044787; PubMed=21843;
RA Beress L., Wunderer G., Wachter E.;
RT "Amino acid sequence of toxin III from Anemonia sulcata.";
RL Hoppe-Seyler's Z. Physiol. Chem. 358:985-988(1977).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=93343891; PubMed=8102051;

```

Db 9 CCCCCVQKR 18

RESULT 7
Q8WLR1 ID Q8WLR1 PRELIMINARY; PRT; 21 AA.
AC Q8WLR1;
AD 01-WAR-2002 (TREMBlrel. 20, Created)
DT DT 01-WAR-2002 (TREMBlrel. 20, Last sequence update)
DT DT 01-WAR-2002 (TREMBlrel. 20, Last annotation update)
DE Major histocompatibility complex class I (Fragment).
GN Name=HLA-B;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Blood;
PC RC
RX MEDLINE=97246738; PubMed=9089103; DOI=10.1007/s002510050227;
RA Vargas-Alarcon G., Gomez-Casado B., Martinez-Laso J., Granados J.,
RA Laysisse Z., Alegre R., Arnaiz-Villena A.;
RT "Differences in intron 2 sequences between B*39061 and B*39062 in
RT Amerindians: comparison with those of B*3901, B*5101, and B*52012
RT alleles."
RL Immunogenetics 45:436-439(1997).
DR EMBL; L77205; AAL40076.1; -.
FT NON TER 1 1
FT NON TER 21 21
SQ SEQUENCE 21 AA; 1972 MW; 5941ECS3ECF3A73D CRC64;

Query Match 0.8%; Score 47; DB 2; Length 21;
Best Local Similarity 56.2%; Pred.No. 2.7e+04;
Matches 9; Conservative 4; Mismatches 1; Indels 2; Gaps

QY 622 NRPPVAV--AGPDKEI 635
||| | : : : : :
DB 4 NRPPAALGGGPRDL 19

RESULT 8
SK1-TITCA STANDARD; PRT; 23 AA.
ID SK1_TITCA
AC F83243;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Toxin Tc1 (Alpha-KTX 13.1).
OS Tityus cambridgei (Amazonian scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butridia; Buthoidae; Tityidae; Tityus.
ON NCBI_TaxID=184226;
RX [1]
RN SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RP TISSUE=Venom;
RC MEDLINE=20565534; PubMed=11113450; DOI=10.1016/S0014-5793(00)02253-5;
RX Batista C.V.F., Gomez-Launas F., Lucas S., Possani L.D.;
RA "Tc1, from Tityus cambridgei, is the first member of a new subfamily
RT of scorpion toxin that blocks K(+)-channels.";
RN FEBS Lett. 486:117-120(2000).
RX [2]
RN STRUCTURE BY NMR.
RP PubMed=11790849; DOI=10.1110/ps.33402;
RX Wang I., Wu S.H., Chang H.K., Shieh R.C., Yu H.M., Chen C.;
RA "Solution structure of a K(+)-channel blocker from the scorpion Tityus
RT cambridgei".
RT Protein Sci. 11:390-400(2002).
CC -1- FUNCTION: Blocks reversibly Shaker B potassium channels. Also
CC displaces binding of noxiustoxin to mouse brain synaptosome
CC membranes.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC CC -1- MASS SPECTROMETRY: MW=2446.4; METHOD=MALDI; RANGE=1-223;

```

RX MEDLINE=22990536; PubMed=14629033;
RA Bull J.J., Badgett M.R., Rokyta D., Molineux I.J.;
RT "Experimental evolution yields hundreds of mutations in a functional
RL viral genome.";
RT J. Mol. Evol. 57:241-248(2003).
DR ENBL; AY264778; AAP34129.1; -.
DR ENBL; AY264777; AAP34075.1; -.
SQ SEQUENCE 29 AA; 3217 MW; 5FF0F978F980B9D1 CRC64;

Query Match 0.8%; Score 46; DB 2; Length 29;
Best Local Similarity 47.1%; Pred. No. 4.7e+04;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 11 LLLLVTTIAGCARKQCSE 27
DB 5 MLLLIIVIGCLALYCSD 21
:||||:|||||
:||||:|||||

RESULT 11
TAT_SIVM2 STANDARD; PRT; 32 AA.
ID AC P05912;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE TAT_Protein (Transactivating regulatory protein) (Fragment).
DE Names:TAT; GN
OS Simian immunodeficiency virus (Mm251 isolate) (SIV-MAC).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11734;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88122665; PubMed=2893293; DOI=10.1038/331619a0;
RA Kestler H.W., Li Y., Naidu Y.M., Butler C.V., Ochs M.F., Jaenel G.,
RA King N.W., Daniel M.D., Desrosiers R.C.;
RT "Comparison of simian immunodeficiency virus isolates.";
RL Nature 331:619-622(1988).
CC -!- FUNCTION: Transcriptional regulator that acts by binding to the
CC trans-activating responsive sequence (TAR) RNA element and
CC activates transcription initiation and/or elongation from the LTR
CC promoter.
CC -!- SUBUNIT: Binds cyclin T1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- MISCELLANEOUS: This is a macaque isolate.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; X06879; -; NOT ANNOTATED_CDS.
CC DR HIV; X06879; TAT$NM251.
CC DR Activator; AIDS; Nuclear protein; RNA-binding;
CC Transcription regulation.
CC FT NON_TER 1
CC SQ SEQUENCE 32 AA; 3556 MW; 211547CBB6F4D817 CRC64;

Query Match 0.8%; Score 46; DB 1; Length 32;
Best Local Similarity 42.1%; Pred. No. 5.3e+04;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 666 PSAVEMENIDKAIATVTGL 684
DB 12 PEKAKKETVEKAVATAPGL 30
| : : : : :
| : : : : :

RESULT 12
Q70222 PRELIMINARY; PRT; 34 AA.
ID Q70222
AC Q70222;

```

Matches 10; Conservative 7; Mismatches 16; Indels 11; Gaps 1

QY 346 TLDPNEVELKAFVAPPVETTYNENWLISHPTDYQGSEIKQH 389
| | : : :
D6 2 TRPNNTRKGVIYGP-----RAVFATDIIGIRQA 34
| | : : :
| | : : :

RESULT 14

Q9YI81 PRELIMINARY; PRT; 27 AA.

ID Q9YI81;
AC Q9YI81;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Engrailed homeodomain protein (Fragment).
OS Priapulus caudatus.
OC Eukaryota; Metazoa; Priapulida; Priapulidae; Priapuluss.
NCBI_TaxID=37621;
RN [1]
RP SEQUENCE FROM N.A.
RA de Rosa R., Grenier J.G., Andreeva T., Cook C.E., Adoutte A., Akam M.,
RA Carroll S.B., Balavoine G.;
RT "Hox genes in Brachiopods and Priapulids: implications for Protostome
RT evolution.";
RL Nature 0:0-0(1999).

-1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC EMBL; AF144895; AAD40651.1; -.
GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeodomain_like.
DR PROSITE; PR00024; HOMEBOX.
DR PRINTS; PS50071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
KW NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3288 MW; 321AE16A0529F351 CRC64;

Query Match 0.8%; Score 45; DB 2; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.7e+04;
Matches 13; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

QY 829 LTEQRKDITLVROLAVLLNVLDSDIKV 854
| | | : : :
D6 6 LTEQRRELAKOLA--LN--ESQIKI 27
| | | : : :
| | : : :

RESULT 15

Q7R7E1 PRELIMINARY; PRT; 29 AA.

ID Q7R7E1;
AC Q7R7E1;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
GN Name=PY07646;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L7XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
RA Cho J.K., Quackenbush J.R., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates P.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,

RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii";
EL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01002846; EAA20148.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 29 AA; 3752 MW; 3497968C69392CCB CRC64;

Query Match 0.8%; Score 45; DB 2; Length 29;
Best Local Similarity 36.0%; Pred. No. 5.2e+04;
Matches 9; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 935 LMWENLIORYIWDGESNCEWSIFVY 959
Db 1 LLLLEYFIKLYIWEYNNNNNNNAFY 25

Search completed: October 12, 2005, 10:30:37
Job time : 69 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	0.9	32	2	S55937	translation initia
2	50.5	0.9	25	2	I40692	cenA protein (IGa
3	49	0.9	25	2	B49111	POU homeodomain pr
4	47.5	0.9	27	1	TZ423	toxin III - snake-
5	44	0.8	34	2	C56635	tubulin alpha chai
6	44	0.8	35	1	NTSRPM	neurotoxin P2 - sc
7	42	0.8	22	2	S58433	isocitrate dehydro
8	41	0.7	25	1	SMWR	metallothionein -
9	41	0.7	29	1	Q1BP57	gene 1.5 protein -
10	41	0.7	35	2	F95172	hypothetical prote
11	40.5	0.7	28	2	C83797	hypothetical prote
12	40.5	0.7	31	2	D84404	hypothetical prote
13	40.5	0.7	31	2	T29628	hypothetical prote
14	40.5	0.7	31	2	G97596	hypothetical prote
15	40.5	0.7	34	2	I67427	retinoic acid rece
16	40	0.7	20	2	S63483	coenzyme F420 hydr
17	40	0.7	28	2	D49829	T-cell receptor va
18	40	0.7	29	2	G39590	neural cell adhesi
19	40	0.7	30	2	S12365	gamma-crystallin -
20	40	0.7	32	2	C49480	major immunophilin
21	40	0.7	33	2	I53186	ISI/cat protein -
22	40	0.7	35	2	JX0200	Gurmarin - Gymnema
23	39.5	0.7	23	2	A48186	ATP synthase beta-
24	39.5	0.7	28	2	S08569	ribosomal protein
25	39.5	0.7	30	2	A49312	cysteine proteinase
26	39.5	0.7	34	2	I67428	retinoic acid rece
27	39	0.7	18	2	PT0286	Ig heavy chain CDR
28	39	0.7	20	2	A60372	pollen allergen Po
29	39	0.7	27	2	PN0583	tyrosine 3-monooxy

Db 2 SVSTPPT---PSPSTP-----PTPSP 19

RESULT 3

B49111

POU homeodomain protein (alternatively spliced) Oct-1B - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: B49111
R:Das, G.; Herr, W.
J. Biol. Chem. 268, 25026-25032, 1993
A:Title: Enhanced activation of the human histone H2B promoter by an Oct-1 variant gene
A:Reference number: A49111; MUID:94043371; PMID:8227066
A:Contents: NTERA 2D1
A:Accession: B49111
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-25 <DAS>
A:Note: sequence extracted from NCBI backbone (NCBIP:139679)

Query Match 0.9%; Score 49; DB 2; Length 25;
Best Local Similarity 34.5%; Pred. No. 3.6e+03;
Matches 10; Conservative 5; Mismatches 10; Indels 4; Gaps 1;

QY 290 SPVLVTPGSTEHSIPTPTPTSAAPSESTP 318

Db 1 SPILAIFPSTSLAQTPP----PTTQP 25

RESULT 4

TZAZ3

toxin III - snake-locks sea anemone
C:Species: Anemonia sulcata (snake-locks sea anemone)
C:Date: 30-Apr-1979 #sequence_revision 24-Sep-1981 #text_change 09-Jul-2004
C:Accession: A91446; A91674; A01798
R:Martinez, G.; Kopeyan, C.; Schweitz, H.; Lazdunski, M.
FEBS Lett. 84, 247-252, 1977
A:Title: Toxin III from Anemonia sulcata: primary structure.
A:Reference number: A91446; MUID:78084776; PMID:23311
A:Accession: A91446
A:Molecule type: protein
A:Residues: 1-27 <MAR>
A:CROSS-references: UNIPROT:P01535
R:Beress, L.; Wunderer, G.; Wächter, E.
Hoppe-Seyler's Z. Physiol. Chem. 359, 985-988, 1977
A:Title: Amino acid sequence of toxin III from Anemonia sulcata.
A:Reference number: A91674; MUID:78044787; PMID:21843
A:Accession: A91674
A:Molecule type: protein
A:Residues: 1-21, 'SC', 24-27 <BER>
A:Comment: Three disulfide bonds are present.
C:Superfamily: toxin III
C:Keywords: venom

Query Match 0.9%; Score 47.5; DB 1; Length 27;
Best Local Similarity 28.1%; Pred. No. 4.9e+03;
Matches 9; Conservative 3; Mismatches 7; Indels 13; Gaps 1;

QY 60 ACCDLSSCDLAWFEGRCYLVSCPKHCEPK 91

Db 2 SCCP-----CYWGGCPWGQNCYPE 20

RESULT 5

C56635

tubulin alpha chain, brain-specific isotype (clone PTU6) - chum salmon (fragment)
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 10-Jul-1998
C:Accession: C56635
R:Coe, I.R.; Munro, R.; Sherwood, N.M.
DNA Seq. 3, 257-262, 1992
A:Title: Isolation of different brain-specific isotypes of alpha-tubulins from chum salmon
A:Reference number: A56635; MUID:93208376; PMID:1296820

A:Contents: brain
A:Accession: C56635
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-34 <COE>
A:Note: sequence extracted from NCBI backbone (NCBIN:128392, NCBIP:128391)
F:28/Binding site: polyglutamate (Glu) (covalent) #status predicted
F:33-34/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicted
F:33-34/Cross-link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicted

Query Match 0.8%; Score 44; DB 2; Length 34;
Best Local Similarity 32.0%; Pred. No. 1e+04;
Matches 8; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 151 EEMSEYXDDYRELEKDLLOPSGKQE 175

Db 6 EDMAALEKDYEEVGVDSEIEGESEE 30

RESULT 6

NFSRPM

neurotoxin P2 - scorpion (Androctonus mauretanicus)
C:Species: Androctonus mauretanicus mauretanicus
C:Date: 15-Oct-1982 #sequence_revision 15-Oct-1982 #text_change 09-Jul-2004
C:Accession: A01758
R:Rosso, J.P.; Rochat, H.
Toxicon 23, 113-125, 1985
A:Title: Characterization of ten proteins from the venom of the Moroccan scorpion Androctonus mauretanicus
A:Reference number: A94318; MUID:85193276; PMID:3992595
A:Accession: A01758
A:Molecule type: protein
A:Residues: 1-35 <ROS>
A:CROSS-references: UNIPROT:P01498
C:Superfamily: scorpion neurotoxin
C:Keywords: neurotoxin; venom
F:1-18, 4-25, 15-30, 19-32/Disulfide bonds: #status predicted

Query Match 0.8%; Score 44; DB 1; Length 35;
Best Local Similarity 30.4%; Pred. No. 1.1e+04;
Matches 7; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

QY 913 CLLKCSGHGHCPLTKRCICSHL 935

Db 15 CATCGGRGKC--VGPQCLCNRI 35

RESULT 7

S58433

isocitrate dehydrogenase (NAD) (EC 1.1.1.41) chain 2 - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 29-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C:Accession: S58433
R:Zeng, Y.; Weiss, C.; Yao, T.T.; Huang, J.; Siconolfi-Baez, L.; Hsu, P.; Rushbrook, J.I.
Biochem. J. 310, 507-516, 1995
A:Title: Isocitrate dehydrogenase from bovine heart: primary structure of subunit 3/4.
A:Reference number: S58432; MUID:95382766; PMID:7654189
A:Accession: S58433
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-22 <ZEN>
A:CROSS-references: UNIPROT:Q7M3F1
C:Superfamily: 3-isopropylmalate dehydrogenase
C:Keywords: oxidoreductase

Query Match 0.8%; Score 42; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 7.2e+03;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 528 PPVANAGPNHTITL 541

Db 7 PPSAKYGARHTVTM 20

RESULT 10

F95172
hypothetical protein SPI481 [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidon, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle, nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95172
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-35 <KUR>
A;Cross-references: UNIPROT:Q9YFPV9; GB:AE005672; PIDN:AAK75575.1; PID:g14972972; GSPDB:B-G; Experimental source: strain TIGR4
C;Genetics:
A;Gene: SPI481

Query Match 0.7%; Score 41; DB 2; Length 35;
Best Local Similarity 42.1%; Pred. No. 1.6e+04;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 293 LTVPSTGSHSIPTTPPSA 311
:: :: :: :: ::
Db 17 VTILEGSNIHVLPMPSYA 35

RESULT 11

C83797
hypothetical protein BH1179 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: C83797
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai, Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: C83797
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-28 <STO>
A;Cross-references: UNIPROT:Q9KD2; GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB048 A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1179

Query Match 0.7%; Score 40.5; DB 2; Length 28;
Best Local Similarity 38.1%; Pred. No. 1.2e+04;
Matches 8; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 957 FYTVLAFTLIVLTGGFTWLC 977
:- :- :- :- :-
Db 9 FALIVLVFILLVIIGA-AWVC 28

RESULT 12

D84404
hypothetical protein Vng2544h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84404
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li; A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: D84404

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:00:24 ; Search time 57 Seconds
(without alignments)
1403.926 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 5580
Sequence: 1 MAPPTGVLSLLLVITAGC.....VSMNGSIRNGASFYSCKDR 1072

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	443	7.9	183	4	US-09-270-767-32456
2	312	5.6	1300	4	US-09-902-540-9932
3	215.5	3.9	827	4	US-09-248-796A-17307
4	206.5	3.7	372	4	US-09-902-540-10204
5	188	3.4	1093	4	US-09-902-540-13786
6	184.5	3.3	3969	3	US-08-061-376-5
7	184.5	3.3	3969	4	US-09-538-092-1262
8	183.5	3.3	822	2	US-08-474-067-9
9	183.5	3.3	822	2	US-08-474-068A-9
10	183.5	3.3	822	2	US-08-472-481-8
11	183.5	3.3	2870	4	US-09-478-467A-15
12	183.5	3.3	3178	4	US-09-479-467A-4
13	183	3.3	2254	4	US-09-949-016-9270
14	182.5	3.3	1020	4	US-09-434-123A-9
15	180	3.2	2736	4	US-09-252-991A-30227
16	179	3.2	880	4	US-09-902-540-10174
17	176.5	3.2	2231	1	US-08-153-799-16
18	176.5	3.2	3913	4	US-09-949-016-10933
19	176.5	3.2	4377	4	US-09-949-016-6978
20	173.5	3.1	1706	2	US-08-453-568-2
21	173.5	3.1	1706	2	US-08-399-411-2
22	173.5	3.1	1706	2	US-08-516-859A-2
23	173.5	3.1	1706	3	US-09-586-472-2
24	173.5	3.1	1706	4	US-09-528-706-2
25	173	3.1	2315	4	US-09-543-681A-5434
26	171.5	3.1	1810	4	US-08-793-273C-4
27	171.5	3.1	1810	5	PCT-US95-11684-4

28 170.5 3.1 1861 2 US-08-790-912-4 Sequence 4, Appli
29 168.5 3.0 4302 3 US-08-658-136-5 Sequence 5, Appli
30 168.5 3.0 4303 2 US-08-460-751-2 Sequence 2, Appli
31 168.5 3.0 4303 4 US-09-479-467A-2 Sequence 2, Appli
32 168 3.0 4302 3 US-09-052-469-8 Sequence 8, Appli
33 168 3.0 4302 4 US-08-422-582-8 Sequence 8, Appli
34 168 3.0 4302 4 US-09-052-262-8 Sequence 8, Appli
35 168 3.0 4339 3 US-09-052-469-6 Sequence 6, Appli
36 168 3.0 4339 4 US-08-422-582-6 Sequence 6, Appli
37 168 3.0 4339 4 US-09-052-262-6 Sequence 6, Appli
38 167.5 3.0 385 4 US-09-902-540-10208 Sequence 10208, A
39 166.5 3.0 991 4 US-09-902-540-11984 Sequence 11984, A
40 166.5 3.0 3892 4 US-09-328-352-5503 Sequence 5503, Ap
41 164.5 2.9 610 4 US-09-538-092-1378 Sequence 1378, Ap
42 163 2.9 1596 4 US-08-978-277A-4 Sequence 4, Appli
43 162.5 2.9 1439 2 US-08-449-644-2 Sequence 2, Appli
44 162.5 2.9 1439 2 US-08-087-244A-2 Sequence 2, Appli
45 162 2.9 1266 3 US-08-506-296B-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-270-767-32456
; Sequence 32456, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32456
; LENGTH: 183
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-32456

Query Match 7.9%; Score 443; DB 4; Length 183;

Best Local Similarity 50.3%; Pred. No. 1.8e-27;
Matches 92; Conservative 22; Mismatches 69; Indels 0; Gaps 0;

Qy 473 INGPFIKETSVDSPVLRLSNLDPGNTSFRITVTDSDGATNTTAAALIVNNAVDPYPVAN 532
Db 1 ISGPFGYQVLPVNTLQDLTSPGNTFKLITVTDSDNNVTNTTATIAVLKETYAPVAN 60
Qy 533 AGPNHTITLPONSITLNGNOSDDHQIVLWLSLGPSEGHVVMQGVQTPYLHLSAMQE 592
Db 61 AGDAVILYLPNNVTLNGTASSDDHEIVAWETWKDASDEAKAVDMQNTTRTPYVQLSNLEE 120
Qy 593 GDYTFQLKVTDSRQOSTAXVTIVQENRPPVAVAGPKELIFFVESATLDGSSSDDD 652
Db 121 GMYTFVLKVTDSGQSGSTAKVHVFPKPTNSPPVAEAGSNTTSLPINWLLNGSDSKDD 180
Qy 653 HGI 655
Db 181 IGI 183

RESULT 2

US-09-902-540-9932
; Sequence 9932, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: *Myxococcus xanthus* Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10 (15849)B

QY	115	DYDMMMLNR--GSPSGIWDGSDPEDIRKDLXFLGKDW-----GLEEMSEYKDDYRELEKDLL	166
DB	217	EMGHYISNRLIGNASGLTNNQGRS-----MGEWADVHALMLMVRGGDDY-----	261
QY	169	QPSGKOEPRGSABYTDWLLPGSEGAFNSSVGDSPAVPAETQODPELHYLNESASTPAPK	228
DB	262	-----ARGA--YAVGVYVGGGGG-----DGYTYGVRRAPATAPT	295
QY	229	L-PERSVLL-----PLPTT-----PSSGEVLEKEKASQOE-----	258
DB	296	LASSRNALTFRHIENGNPULTSHPLAYGQAGLGNNAVHSSGEW-----ASMLUECVSVLL	351
QY	259	-----QSSNSSGKEVLMPSHSLPPASLELSSVTVKSPVLTVTPGSTSEHSIPTPTSAAP	313
DB	352	QAHPPQEAQDRMKHYLVAGYKATFVS---PTFLDARDALLATAAS-----	394
QY	314	SESTPSELPISTTAPRTVKELTVSAGDNLIITLPDNE-----VELKAFVAPAPPVETTVN	369
DB	395	-----DPADYQRFLOAF-AKRGAGMGAKAPDRSDMDHGVGVSPAGNFIEVT--S	442
QY	370	YEWNLISHTDVOGEIKQGHKQTLNLSQSVG---LYVPKVTVSS-----	412
DB	443	LRDDDALECGDKDGLVDGETGLLRVSVRNVGSLGAFSGTVSTTSATLFFPSGNTL	502
QY	413	--NAPGEGFVNTVKPARVNLPPVAVVSPQL--QELTLPLTSAL-----IDGS	457
DB	503	TFSSMGRGGTATATLPLVRLAAQPNNAHGLRITTFDEPSLPASASSTTFDPVNVDEVPGA	562
QY	458	QSTDDTEIVSYHWE-----BING-----PFIEEKTSDVSPVLRLSN	493
DB	563	AAEDDEGALSSWTSSTLWNPYGEADWKVQGPPEGGRYMHAPNPAVQADLFLISPMKVNE	622
QY	494	LDPGVSPRLTVTDS-----DGATNSTT-----AALIWNNAVDPVPA	531
DB	623	TGAFSPSFRYHSFESDIYPGGSIFPYDGAVLEFTHDGLWYDIAAMGVASIIYSIDPTD	682
QY	532	NA-----QPNHTITLPQNSITLNGQSDDDHQIVLY-EWS-----LGPSSGKHVVMQ-	578
DB	683	GSPGYTEQGN---PLMYRNAMVQGVQNTQENKTVRFPPEWSTAEURLADMFAGKNVQFRF	739
QY	579	--GVQTPYLHLSAMQGDYTFOL---KVTDSRQSQSTAXTVTVIOPEN---NRPPVAVAG	630
DB	740	RIGTDYAV-----GAYGFDLDDVRFNVNRMPPSGRVD--EQGDGATCQNPVAVAG	789
QY	631	PD-----KELIPFVESA-----TLDGSSSSDDHG-IVFYHWEHVRGSPSAVEMENI	674
DB	790	RDVGSPSPAPYVYRVRNPETGALELVRLHLDGTASVDPEGOPLTYHTWTVSGP-AVTLVNA	848
QY	675	DKAIAVTG--LOVGYTHFRLTVKQOGLSSSTSTLTVAVKKNSSPPRARACGRHVVLVP	732
DB	849	DTATPSFTADVDDSDSILTEQLVVRDGLTSAHOVRAWLVNVRNAPTAVVVTAPEAEVFS	908
QY	733	NNSITLDGSRSTD-QDRIYSVYLWIRGQSPAGDVIDGSDHSDVALQTLNL-VEGVVTFHL	790
DB	909	SEPVTLIDARASDPDGLISYKWAQT-SGPAVA-LKAGSSATPSFKVPEVAVDTQFTFL	966
QY	791	RVTDSQASDSDTATVEVQ	809
DB	967	VVNHCRDQSKTATVTVVVR	985
<p>RESULT 6</p> <p>US-08-061-376-5</p> <p>; Sequence 5, Application US/08061376</p> <p>; Patent No. 6175000</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Evans, Glen A.</p> <p>; APPLICANT: Djabali, Malek</p> <p>; APPLICANT: Selleri, Licia</p> <p>; APPLICANT: Parry, Pauline</p> <p>; TITLE OF INVENTION: CHARACTERIZATION OF A CHROMOSOME 110233</p> <p>; TITLE OF INVENTION: TRANSLOCATION BREAKPOINT ASSOCIATED WITH ACUTE LEUKEMIA</p>			

1	NUMBER OF SEQUENCES: 12	Query Match	3.3%; Score 184.5; DB 3; Length 3969;	
2	CORRESPONDENCE ADDRESS:	Best Local Similarity	20.2%; Pred. No. 0.00016;	
3	ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark	Matches 181; Conservative 115; Mismatches 273; Indels 329; Gaps 47;		
4	STREET: 444 South Flower Street, Suite 2000			
5	CITY: Los Angeles			
6	STATE: California			
7	COUNTRY: USA			
8	ZIP: 90071			
9	COMPUTER READABLE FORM:			
10	MEDIUM TYPE: Floppy disk			
11	COMPUTER: IBM PC compatible			
12	OPERATING SYSTEM: PC-DOS/MS-DOS			
13	SOFTWARE: Patent In Release #1.0, Version #1.25			
14	CURRENT APPLICATION DATA:			
15	APPLICATION NUMBER: US/08/061.376			
16	FILING DATE: 13-MAY-1993			
17	CLASSIFICATION: 435			
18	ATTORNEY/AGENT INFORMATION:			
19	NAME: Reiter, Stephen E.			
20	REGISTRATION NUMBER: 31,192			
21	REFERENCE/DOCKET NUMBER: P41 9387			
22	TELECOMMUNICATION INFORMATION:			
23	TELEPHONE: (619)546-4737			
24	TELEFAX: (619)546-9392			
25	INFORMATION FOR SEQ ID NO: 5:			
26	SEQUENCE CHARACTERISTICS:			
27	LENGTH: 3969 amino acids			
28	TYPE: amino acid			
29	STRANDEDNESS: unknown			
30	TOPOLOGY: unknown			
31	MOLECULE TYPE: protein			
32	US-08-061-376-5			
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				

Db 3057 SQISNAVQTPPHLPKPAATEKLIIVNQNMQLVYLQPLNGVTKIQITQSVSTPSVME 3116
Qy 367 TY-----NYENNLISHPTDYQGEIKQGHKQTL-----N 394
Db 3117 TNTSVLPGMGGLTLTGLNPSLPTQSGLFSPASGLLPMSSHQHLHSPFAATQSSPPPN 3176
Qy 395 LSQLSVGLYV-----FKVTVSSENAFEGFVNVT-----KPARRV----- 430
Db 3177 ISNPPSGLLIGVQPPDPQLLVSESSQRT-----DLSTTVATPSSGLKRPISRLQTRKNK 3232
Qy 431 -----NLPPVAVSPQ-LQELT---LPLTSALID-GSQSDTDDTIVSYHWEINGP 476
Db 3233 KLAPSSPTSNTAPSDVSNWTLINFTPSQLPNHPSLLDGLSLNTSSHRTV-----P 3283
Qy 477 FI--EETSV-----DSPVL-----RLSNLDPGNYSFRLTVDSDGA 511
Db 3284 NIKRKSIMYFEPAPLLPQSVGTAATAAGTSTISQDTSHLTSGSVS-----GL 3334
Qy 512 TNSTTAALIVNADVPPVANAG-PNH-TITLPO-----NSITLNGQSS-----DD 556
Db 3335 ASSSVLVNVMQTTTPTSSASVFGHVTLTNPRLLGTPDIGSISNLLIKASQOQSLGIQD 3394
Qy 557 HQIVLYEWSLGPGEKGHVMOGVQTPVYLHLSAMQEGDYTFQLKVTDSRRQGSTAXVTVI 616
Db 3395 QPVAL-----PPSGMPFQLGTSQP-----STAAITAAASIC 3427
Qy 617 VQPENRPPVAVAGDKELIPFVESATLDGSSDDHGIVFYHWEHVRG--PSAVEMENI 674
Db 3428 VLPSTQTGTITAAAP-----SGEADEH---YQLQHVNQLLASKTGIHSS 3468
Qy 675 DKAIATVTLGVTHFRLTVKDOGLSSTLTVAVKKNNSPPRARAGRHVVLVLPNN 734
Db 3469 QRDLDSAGPOVSFNFTQTVDAFNSMGLRQKALSSAVQASPTSP-----CG-----SP 3516
Qy 735 SITLDGSRSTDDQRIVSYLWIRDGOSPAAG-----DVIDGSDHSHVALQL 778
Db 3517 SSPSGORS-----SPSVPGFTKPKTKRFPQLPLDKNGKKGKVS--- 3558
Qy 779 TNLVEGVYTFHLRVTDGQ-----ASDITDITAVEVQDPDRKSG 816
Db 3559 -----HLRTSSSEAHIPDQETTSLSGTGTPGAEEQDQTSVE-QSSQKEG 3605

RESULT 8

US-08-474-067-9
; Sequence 9, Application US/08474067
; Patent No. 5811518
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,067
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/607,293

; FILING DATE: 30-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1682
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-474-067-9

Query Match 3.3%; Score 183.5; DB 2; Length 822;
Best Local Similarity 21.4%; Pred. No. 1.4e-05;
Matches 165; Conservative 90; Mismatches 260; Indels 257; Gaps 40;

Qy 359 APAPPVETTYNYE-----WNLSHPTDYQGEIK-----QGHKQTLNLSQLSVGLYFKVTVSS 411
Db 142 ADSPP-EGVFTIEKESGMLLLHMLDREKIVKELYLYGHAVSENGASVEEPMNI-SIIIVTD 199
Qy 412 ENAFGEFVNVTVKPARRVNLPPVAVVSPOLQELTLPLTSALIDGSDTDD-----TEIV 466
Db 200 QDNKPKFTQDTRGS-----VIEGMFGTSMQVMTATDEDDAVNTYNGVV 245
Qy 467 SY--HWEINGPF-----IBEKTSVDSPLRLSNLDPGNY-SFRLTV--TDSGATNSTT 516
Db 246 AYSIHSQEPKEPHDLMTIHKSTGTISVI--SSGLDREKVPYRLTVQATDMDCGEGSTT 303
Qy 517 AALIV-----NNADVPPVANAGPNHTITLQNSITLNGQSSDDHQIVLYEWSLGPGS 570
Db 304 AEAVVQIILDANDNAPEFEP-----QKYEAWVPENEV-----GH 336
Qy 571 EGKHVVMQGVQTP-----YLHLSAMQEGDYTFQLKVTDSRRQGSTAXVTVIQPENNR 623
Db 337 EVQRLTVDLDVFNWPNARATYTHIVGGDDGQH-----FTTTHPETNQ 379
Qy 624 PPVAVAGPDKELIPFVESATLDGSSDDHGIVFYHWEHVRGPGSAVEMENIDKAIATVTG 683
Db 380 ---GVLTKKGLDF-----EAQDQH--TLYVEVTNEAPFAVKLP----- 413
Qy 684 LQVGYTHFRLTVKDOGLSSTLTVAVKKNNSPPRARAGRHVVLVLPNNISITLDGSR 743
Db 414 -----TATATVVHVVKOVNEAP-----VFVPPSKVIEAQEGIS 446
Qy 744 -----TDDQRIVSYLWIRD-----GQSPAAGDVIDGSDHSHVALQL 778
Db 447 IGEIVCIYTAQDPDKEDQKI-SYITISRDPAWLAVIDPDSGQITAAAG-ILDRDSQF----- 500
Qy 779 TNLVEGVYTFHLRVTDGQASDITDITAVEVQDPDRKSGELVELTLQVGVGQLTEQRKQTLV 838
Db 501 --VKNNVYEVWVLATDSGNPTTGTGL-----LLTLTDINDHGPIPEPRQIIIC 548
Qy 839 RQLAV--LLNLVDSDIKVKIRAHSDLSITVIVVQSRPPPKVLKAAFAVARNLHMRUSKE 896
Db 549 NQSPVPQVLNITDKDLSFNS-----SPFOALQTHSDSDIYMAEVS-----E 589
Qy 897 KADFL-----LPKVLAVDTAGCLLKCSGHCHCDPLT--KRCICS-HLWMENLQRIWDGES 950
Db 590 KGDTVALSLKKFLKQDQTYDLHLSLSDHGNREQLTMRATVCDCHGQVFNDCPR-PWKGG- 647
Qy 951 NCEWSIFVTVLAFTLIVLGGFTWLCTCCCKRQKTKIRKTKYITLIDNMDQERMELR 1010
Db 648 -----FILPILGAVLALT---LLALLLVKKR-----KVKEPLLLPEDD----- 686
Qy 1011 KYGIGKHRSTHNSLMVSESEFSDQDTIFPREKMERG-----NPKVSMNGSI 1059
Db 687 -----TRDNVFFYYGEGGGBEDQD--YDITQLHRGLEARPEVILRNDV 727

RESULT 9


```

; PRIOR APPLICATION NUMBER: 60/115,127
; PRIOR FILING DATE: 1999-01-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 3178
; TYPE: PRT
; ORGANISM: C. Elegans Lov-1 protein
US-09-479-467A-4

```

```

Query Match      3.3%; Score 183.5; DB 4; Length 3178;
Best Local Similarity 19.4%; Pred. No. 0.00013;
Matches 147; Conservative 97; Mismatches 259; Indels 253; Gaps 29;

```

```

QY 195 FNSVGDSPAVPAETQQOQDELHYLNESASTAPKLPERSVLLPLPTTPSSGEVL-----248
DB 327 FTSTASTSTSSITQQ-----SSSITSPSSSTLSTSIPTTTTPEITSTLSSLPDN 379
QY 249 -----EKEKASQLOEQSSNSGKEVLMPSHSL-----PP 277
DB 380 AICSVLDETTSTTTTMTLSTTTEEPSTSTTTTEVTSTSTVTTTEPTTLTTSTAST 439
QY 278 ASLELSSVTVKSP-----VLTVPGSTEHSIPTPTSAAPSESTPSELPISTTAP---329
DB 440 STTEPSTVTTSPSTSPVTSTVSSSSSTVTTPTSTESTSTSPSTVTTSTTAPSTS 499
QY 330 -----RTVKELTVSAGDNLIIITLDPNEVE 353
DB 500 TTGPPSSSTSPSTASSVSSSTASTQSTSTQOQSTTTKSTTTSSDG---TNPDFV 556
QY 354 LKAFVA--PAPPVETTYNEMNLISHT-----DYQGEIKQGHKQTLNLSQ-----397
DB 557 EKATFTFYDSTSVNLTLNSGLIGVQTSIECTSPSTSSNYVSTTKDGCFTKSVSMPLG 616
QY 398 -----LSVGLYVFKTVSSE-----NAGGEGFVNTVVPARRVNLPPVAVS- 439
DB 617 GTPASTFVGPNGYTFRAITMTDDKKVYVYANVYIOEYSSTIESSEST-----AVASS 672
QY 440 ----PQLQELTLPLTSALIDGS--OSTDDTEIVSVHWBEI--NGPFIIEKTSVDSPVLRLS 492
DB 673 TSTSTSPSTLS--TSVTPESTSTSSDSTTSAGSTTTLQESTTTSESTDSSTTIS 731
QY 493 NLDPGNYFRLTVDSDGATNSTAALIVNNAVDYPPVA--NAGPNHTITLQNSITLN--549
DB 732 DTSTSTSPSTADSTSTLSVDQDFILDGSLSMWNETRHEDSINIVPLPTNAITPTER 791
QY 550 -----GNOSDDHQIVL-----YEWSLGPGSEKHHVMQGVQTPYLHLSAMQGDYTF 597
DB 792 SQTFCRNVTSTPPLIKESTCLNYSNTVLNATYSNIPQIETFLVGI-----GTVEF 846
QY 598 QLKVTDSRQQTAKVTVIVQPNRNPPVAVAGPKELIFPVESATLDGSSSSDDHGIVF 657
DB 847 RINMTDLTMQVSHIFTL-----NVVA--DSTSTSEVTSTTSGSSE-----888
QY 658 YHWEHVRGSPAVEMENIDKAIATVTGLQVGYHFLRTLVKDOQGLSTSTLTAV--KKEN 715
DB 889 -----SSAISTTSGIE-----STSTLEASTTSDASQD 914
QY 716 NSPPRAGGRHVLVLPNNSITLDGRSTDDORIVSYLWIRDOQSPAAGDVIDGSDHVA 775
DB 915 SSTSTSDSG-----TTSSTTIDSNST-----PSTSD-----SSG 945
QY 776 LQLTNLVEGVYTHFLRVTDSCASDITDTATVEVQPD 811
DB 946 LSQT-----PSDSSASDUS-MRTTVDPD 968

```

```

RESULT 13
US-09-949-016-9270
; Sequence 9270, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USRS THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9270
; LENGTH: 2254
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9270

```

```

Query Match      3.3%; Score 183; DB 4; Length 2254;
Best Local Similarity 20.8%; Pred. No. 8.2e-05;
Matches 174; Conservative 83; Mismatches 302; Indels 276; Gaps 37;
QY 170 PSGQOEPRGS-----AEYTDNGLLPGS-----EGAFNS-----SVGDSPAVP 206
DB 1325 PERKEEPPRPFLBQPLLGELTVGTVPDLSURLSWTVAQGFDFSMVQYKDAQGQFQAVP 1384
QY 207 AETQ-----DPELHY-----LNESAS-----TPA 226
DB 1385 VAGENETVPGLPDDRKIKNLYGLGRQVRGPEVVAKTAPQEDVDDETSTPBLGHEA 1444
QY 227 PKLPERSVLLPLTPSPSSGEVLE-----KEKASQLOEQSSNSGKEVLM 271
DB 1445 PESPEPLLGELTVTGSSPDSLSLFTVPQGSFDSFTVQYKDRGRPRAVRVGKSEVT 1504
QY 272 SHSLPPA-----SLELSSVTVKSPVLTVTPGSTEHSIPTPTSAAPSESTPSELPI 323
DB 1505 VGLGEPGHKYMHLGLHEGQVRGVPVAVGTAPQOEE-----TPPATESPLEPRIGELTV 1560
QY 324 SPTTAPRTVKELTVSAG--DNLIITLPDNEVELKAFVAPAPVE--TTYNYEWNLSHPTD 380
DB 1561 TDVTNSVGLSWTVPEGQFDSFIQYKDKQGVQVVAADQOREVTVVNL-----1611
QY 381 YQGEIKQGHKQTLNLSQSLVGLYVFKVTSSENAFAFGEFVNVTVKPARVNLPPVAVVSP 440
DB 1612 -----PERKYMNYGLHDG-----QRMGLSVVITVAPATEASKPP---LEP 1651
QY 441 QLQELTLPLTSALIDGOSTDDT--EIVSYHWEERINGPFIEKTSVDSPVLRLSLNDPGNY 499
DB 1652 RLGLTV-----TDITPDSVGLSWTVPEGEF-----D 1678
QY 500 SFRLTVDSDGATNSTTAALIVNNAVDYPPVANAGPNHTITLQNSITLNGNOSSDDHQI 559
DB 1679 SFVQYKDRDQOQ-----VVPVAAD-----QREVTIPDLERKRYK 1716
QY 560 VLYEWSLG-----PGSEGHVVMQGVQTPYLHLSAMQEGDYTFOLKVTDSRQQTAXVTV 615
DB 1717 LLFGIQDGKRSPSVSVEAKTVARGDASPG---APPLG---ELWVTOPTDPSLRLSWTV 1769
QY 616 IQVQENRPPVAVAGPKELIFPVESATLDGSSSSDDHGIVFYHWEHVRGSPSAVEMENID 675
DB 1770 ---PEGQFDSFVQFQKDX-----GPQVFPVEGHE 1796
QY 676 KAIATVTGLQVG--TVHREL-----TVKDOQGLSSTSTLTVAVKKENNSP 718
DB 1797 RSV--TVTPLDAGRKYRFLLYGLLKKRKGPLTADGTTTEARAMDITGT-----KRP 1846
QY 719 PRARAGGR--HVLVLPNNSITLDGS--RSTDDQRIVSYLWIRDOQS---PAAGDV-----ID 768
DB 1847 PKPRLGEBELQVTTVTQNSVGLSWTVPEQGFDSFVQYK--DRDQGVQVVPVEGSLREVSV 1905
QY 769 GSD--HSVALQITNLVEGVYTHFL--RVITDSQASDITDTATVEVQPDPRKSLVELTLQVG 825

```

Db 1906 GLDPAHRYKLLYGLHHCKRVGPISAVAITAGRETETETATTPTTAPPEPHGLBELTVEEA 1965
Qy 826 VG-----QUTERQKOTLVRLQVALLNVLDSIDKVKIRAHSDLSLTVIVYQOS 873
Db 1966 TSHTLHLSWMVTEGEFDSFEIQYT-----DRDQLQWVRIGGRNDITLSGLS 2014
RESULT 14
US-09-434-123A-9
; Sequence 9, Application US/09434123A
; Patent No. 6586580
; GENERAL INFORMATION:
; APPLICANT: LINDAHL, Gunnar
; APPLICANT: STAHAMMAR-CARLEWALM, Margaretha
; APPLICANT: STENBERG, Lars
; TITLE OF INVENTION: PROTEIN RIB, A CELL SURFACE PROTEIN THAT CONFERS IMMUNITY TO MANY
; TITLE OF INVENTION: OF THE GROUP B STREPTOCOCCUS: PROCESS FOR PURIFICATION OF THE PR
; TITLE OF INVENTION: KIT AND PHARMACEUTICAL COMPOSITION
; FILE REFERENCE: 0552-0140P
; CURRENT APPLICATION NUMBER: US/09/434,123A
; CURRENT FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Group B Streptococcus - Strain A909
US-09-434-123A-9
Query Match 3.3%; Score 182.5; DB 4; Length 1020;
Best Local Similarity 22.4%; Pred. No. 2.4e-05;
Matches 167; Conservative 83; Mismatches 307; Indels 189; Gaps 38;
Qy 164 EKDLLQPSGKQEB--PRGS-----ABYTDWGLLPGSEGAFNSSVGDSP--AVPAETQQDPEL 215
Db 229 DKDKYDPTGGTTVPQGTTPVDSKEDTDLVKIPDGSKGVPTVVGDRPDNTNPGDHVATVEV 288
Qy 216 HYLNESAS-----TPAPKLPRSVLLPL--PTTFSSGEVLEKEKASQLQSGNSSG 265
Db 289 TYPDGTKDTVEVTVHVTPKP--VPDKDYDPTGGTTVPQGTTPVDSKEDTDLVKIPDGSKG 347
Qy 266 KEVLM--PSSHLPASLELSVTVKESPVLTVPSTGSEHSIPTPTSAAPSESTPSBPLP 322
Db 348 VPTVVGDRPDNTNPGDHVATVEVTPDGTKDTVE--VTVHVTPKPVPDKDYDPTGGTT 405
Qy 323 ISPTTAPRTVKELT-----VSAGD-----NLITLTPD---NEVEL 354
Db 406 V-PQGTVPDSKEDTDLVKIPDGSKGVPTVVGDRPDNTNPGDHVATVEVTPDGTKDTVEV 464
Qy 355 KAFVAPAP-PVETTYNTYENWNLISHPTDYQGEIKQG-----HKQTLNLSQLSVGLYVFKVTV 409
Db 465 TVHVTPKPVPDKDYD-----PTGGETTVPGQTPVDSKEDTDLVKIPDGSKGVPTV 516
Qy 410 SS-----ENAFGE-----PVNTVTKPARRVNI,PPVA--VVSPOQLTLP 448
Db 517 GDRPDNTNPGDHVATVEVTPDGTKDTVEVTVHVTPKP-----VPDKDYDPTGGTTVP 571
Qy 449 LTSALIDGSOSTDTEIVSVHWEINGPFIETKTSVDSPLRLNLNLPNGYSFRLTVDTS 508
Db 572 QGTPVSD-KEITDLVKIP-----DQSKGVPTVVGDRPDNTN-----PGDHVATVEVTP 619
Qy 509 DGATNSTAALI-----VNAVDPYPVANAGPNHTITLPQNSITLNGNSSDDHQLVLY 562
Db 620 DG-TKDTVEVTVHVTPKPVPDKDYDPTGG-----ETTVQ-----GTPVSDKEITDLV 667
Qy 563 EWSLGPSEGKHVM-----QGVQTPYLHLSAMQEGDYFQKLYTDSRSQQSTAXTVIVQ 619
Db 668 K--IPDGSKGVPTVVGDRPDNTN-----GDHVAATVEVTPDGTKDTVEVTVHVT 716
Qy 620 ENNRPPVAVAGPKDELIFPV--ESATLDGSSSDHGHVVFVHWEHVRGPGSAVEMENDKA 677
Db 717 K-----PV-----PDKDYDPTGGTTVPQGTTPVDSK-----EITDLVKIPDGSKG 757

Qy 678 IATVTGLQ-----VGYTHFLTYKQQQLSSTSLTIVAV-----KKENNSPPPARACGR 726
Db 758 VPTVVGDRPDNTNPGDHVATVEVTPDGTKDTVEVTVHVTPKPVPDKDYD-----TGG- 812
Qy 727 HVLVLPNNSITLDGSRSTDDQRIVSYLWIRGQSPAAAGDVIDGSDHSHVALQLTLNLVEGVY 786
Db 813 -----EITVPQGTTPVSDKEITDLVKIPDGSKGVPTVVGDRPD-----TN-VPGDH 856
Qy 787 TFHLRVTDGASDSTDATVEVQDDP 812
Db 857 VATVEVTPDGTKDTVEVTVHVTPKP 882
RESULT 15
US-09-252-991A-30227
; Sequence 30227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30227
; LENGTH: 2736
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30227
Query Match 3.2%; Score 180; DB 4; Length 2736;
Best Local Similarity 20.7%; Pred. No. 0.0002;
Matches 166; Conservative 90; Mismatches 282; Indels 264; Gaps 41;
Qy 171 SGKQEPGRGSAEYTD-----WGLLPGSEGAFNSSVGDSPVAPPAETQQDPE 214
Db 1033 SGTAEPGATVLTLDGSGNPIGQVTDGSGNWSFTP-----STPLADGTVVNA-TATDPA 1085
Qy 215 LHYLNESASTP--APKLP-----ERSVLLPLPTTPPSGSEVLEKEKASQLQSGNSSGKE 267
Db 1086 GNTGGQGTTVDAIAPATPTVNLNNGSSLSGTAEFGTIVLTGNGNPIAEVTADGSGNW 1145
Qy 268 VLMP-----HSLPPASLELSVTVKESPVL-----TVTPGSTHSI 304
Db 1146 TYTPTPIANGTVVNVVVAEDAAGNSPPATVTVDS-SAPPAPVINPNSGVVISGTAEAGA 1204
Qy 305 PTPPTSAPPS-----ESTPSSELPISPTTAPRTVKELTVSAGONLIITLDPNE 351
Db 1205 TVLTLDAGNPIGQVTDGSGNWSFTPTGPANGTVIVATATDPTGNTG-----PQAA 1257
Qy 352 VELKAFVAPAPPVETTYNTYENWNLISHPTDYQGEIKQGHKOTLNLSQLSVGLYVFKVTVSS 411
Db 1258 TVVDAAVAPPAPVIDP-----SNGTTISGTAEAGAKVIL-----TDGN 1294
Qy 412 ENAFGE-----GFVNTVTKPARRVNI,PPVA--VVSPOQLTLP 461
Db 1295 GNPGETTADGSGNWTFTPA-----TFLANGTVVNAVAQD---PAGNTGPGSTTVDAVA 1346
Qy 462 -DTEIVSVHWEINGPFIETKTSVDSPL-LRLNLDP-----GNYSFR----- 502
Db 1347 PNTFVN-----PSNGNLLNGTAEPGSIIVLTLDGNGNPIGQTTADGSGNWSFTPGSQLPNG 1402
Qy 503 --LTVTDSGATN-STTAALIVNAVNDYPP-----VANAGNHTIT----LPQ 543
Db 1403 TVNVNTASDAAGNTSAPATTTVDSSLPISIQVDPNSGNSVIGTADAGNTIIITDGNNGPI 1462
Qy 544 NSITLNGQSSDDHQIVLVYEWSLGPGSEGHVVGQVQTPYLHLSAMQEGDYTQLKVT 603

Db 1463 GQVTADGSGN-----WSFTFG-----IPLPDGTVVNVVARSP 1494
QY 604 SSRQOSTAXVTV-----IVQENRRPPVAVAGPKELIF-----PVESATLDGSSS 649
Db 1495 SNVDSAPAVITVDGVAPAPVIDPSNGTEISGTAEAGATVILTDGGGNPIGOATADGSG- 1553
QY 650 SDDHGIVFYHMEHVRGSAVEMENIDKAIATVTGLQVGYHFRFTVKDQOG-LSSTSTLT 708
Db 1554 -----NWTFTFG-----TPLANGTV-INAVAQDPAGNTSGPASVT 1587
QY 709 VAVKKENNSPPPARAGGRHVLVLPNNSITLDGSRSTDDQRIVSYLMWRDQSPAGDV-I 767
Db 1588 V-----DAIAPAP-----VINPSNGVVISGTAEAGATVILT-----DGNGNPICQVTA 1631
QY 768 DGS DSHVALQLTNLVEGVYTHLRVTDSQASDITD-ATVE-VQP-----DPRKSGLVEL 820
Db 1632 DSGGKWAFTPATPLANGT-VINALAQDAAGNNSPTSATVDSLAPAPVIDPSNGSVIAG 1690
QY 821 TLQVG-----VGQLT 830
Db 1691 TAEAGATVILTDGNGNPICQVT 1712

Search completed: October 12, 2005, 10:19:17
Job time : 62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2005, 10:00:24 ; Search time 52 Seconds
(without alignments)
1983.546 Million cell updates/sec

Title: US-10-764-390-3
Perfect score: 5580
Sequence: 1 MAPPTGVLSLLLVTTIAGC.....VSMNGSIRNGASFYCSKDR 1072
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	4.0	26926	1 I38344	titin, cardiac mus
2	208.5	3.7	6805	2 S20901	titin - rabbit (fr
3	208	3.7	5291	2 F90696	hypothetical prote
4	205	3.7	1742	2 S24600	projectin - fruit
5	203.5	3.6	5188	2 B85547	probable RTX fami
6	202.5	3.6	6658	2 T13931	projectin - fruit
7	201.5	3.6	3283	2 AC1018	large repetitive p
8	200.5	3.6	849	2 D82510	chitinase VCA0027
9	200.5	3.6	1009	2 C64483	hypothetical prote
10	195.5	3.5	1020	2 A46405	C protein alpha an
11	195.5	3.5	2232	2 T34434	hypothetical prote
12	192.5	3.4	1000	2 I46521	titin - rabbit (fr
13	190.5	3.4	1461	2 E90696	hypothetical prote
14	190.5	3.4	1461	2 A85547	hypothetical prote
15	190	3.4	3083	2 AH2493	hypothetical prote
16	188	3.4	528	2 I47141	gastric mucin (clo
17	188	3.4	4006	2 T09070	probable tenascin
18	187.5	3.4	3507	2 T34513	hypothetical prote
19	187	3.4	2013	2 A11489	probable peptidogl
20	187	3.4	3624	2 AD0835	large repetitive p
21	186	3.3	2944	2 A54849	collagen alpha 1(V
22	184.5	3.3	822	1 IUMSCP	P-cadherin precurs
23	183	3.3	6831	2 A88852	protein unc-22 (im
24	183	3.3	6839	2 S57242	twichin [similar
25	183	3.3	7160	2 T27935	hypothetical prote
26	182.5	3.3	1275	2 T33369	hypothetical prote
27	181.5	3.3	3968	2 A44265	trithorax homolog
28	179	3.2	5175	2 T20992	hypothetical prote
29	179	3.2	5198	2 T43290	hemikenttin precurs

RESULT 1

I38344
titin, cardiac muscle [validated] - human
N:Alternate names: connectin
N:Contains: serine/threonine-specific protein kinase (EC 2.7.1.1.-)
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: I38344; I38345; S20897; S20898; S20899; S63665; S37393
R:Labeit, S.; Kolmerer, B.
Science 270, 293-296, 1995
A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
A:Reference number: A57430; MUID:96026330; PMID:7569978
A:Accession: I38344
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A:Molecule type: mRNA
A:Residues: 1-26926 <LAB1>
A:Cross-references: UNIPROT:Q10466; EMBL:X90568; NID:gl017424; PID:gl017425
R:Musco, G.; Tziatzios, C.; Schuck, P.; Pastore, A.
Biochemistry 34, 553-561, 1995
A:Title: Dissecting titin into its structural motifs: identification of an alpha-helix m
A:Reference number: I38345; MUID:95119041; PMID:7819249
A:Accession: I38345
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1977-2014 <MUS>
A:Cross-references: EMBL:X83270; NID:g602579; PIDN:CAA58243.1; PID:g602580
R:Labeit, S.; Gaucel, M.; Lakey, A.; Trinick, J.
EMBO J. 11, 1711-1716, 1992
A:Title: Towards a molecular understanding of titin.
A:Reference number: S20897; MUID:92258380; PMID:1582406
A:Accession: S20898
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 13597-14200, 'I', 14202-14696 <LAB2>
A:Cross-references: EMBL:X64698; NID:g37192; PIDN:CAA45939.1; PID:g37193
A:Accession: S20897
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 16330-16382, 'S', 16384-16756, 'F', 16758-16860 <LAB3>
A:Cross-references: EMBL:X64699; NID:g37190; PIDN:CAA45940.1; PID:g37191
A:Accession: S20899
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 'P', 22278-22431, 'R', 22433-22448, 'G', 22450-22453, 'Q', 22455-22480, 'TR', 22483-2
A:Cross-references: EMBL:X64697; NID:g37190; PIDN:CAA45938.1; PID:g37195
R:Kolmerer, B.; Olivieri, N.; Witt, C.C.; Herrmann, B.G.; Labeit, S.
J. Mol. Biol. 256, 556-563, 1996
A:Title: Genomic organization of M line titin and its tissue-specific expression in two
A:Reference number: S63665; MUID:96177761; PMID:8604138
A:Accession: S63665
A:Status: nucleic acid sequence not shown

Db 14084 SNVT--KNTATVSWKRPVDGSGSEITGYHVRREKKSLRWRAIKTPVSDLRCKVTLQGE 1414
 QY 497 GN-YSPRLTIVTD--SDGATNSTTAALIVNNAVDPYPVANAGPNHTIITLPNSITLNGNQS 553
 Db 14142 GSIFYFRVSNRAGIGPPEASDSVLMDAA-YPGPSPNPVHTDTTKKSASLAWGKPH 1420
 QY 554 SD-----DHQIVLYE-W-----SLGPGSEG 572
 Db 14201 YDGGLEITGVVHEHQVKVGDEAMIKDTTGTALRITQFVPDLQTKEKNFRISAINDAGVG 1426
 QY 573 KHVVMQGVQ-----TPYLHLA-----DYTF---QLKV 601
 Db 14261 EPAVIPDVEIVEREMAPDFELDAELRTLVRAGLSIRIFPIPKGRPAPEVTWTKONIL 1432
 QY 602 TDSRQOSTAKVTVIQEPENRRP-----VAVAGPOKELIFFVESATLDGSSSSDD---- 652
 Db 14321 KNRAIENTESFTLLIIPECNRVDTGKFVMTIENPACKSGFNVRVLDLTARSPQLRPT 1438
 QY 653 ---HGIVFYHWE--HYRGPSAVEMENIDKAIAIAT-----VTGLQVG-TY 685
 Db 14381 DITKDSVTLHWDLPLIDGGSRIINYIEKREATKSYSTATTKCHKCTYKVTGLSEGCEY 1444
 QY 690 HFLRTVKDOOGLSSTLTIVAVKENNSPPRARGRHVLVLNNISLTLDGSRSTDQRI 749
 Db 14441 FPRVMAENEYIGEPTETTEPVKASBASP-----PDLSNIMDITKST---- 1448
 QY 750 VSVLWLR--DGQS PAAGDIVD---GSD---HSVALQ-----LTNLVEG-VTFHLRV 793
 Db 14484 VSLAWPKKHGGSKITGVIEAQKGSDOWTHITTVKGLECVVRNLTEGEEYTFQMVA 1454
 QY 794 DSOGASDTPATVVEVOP--DPRKSGVLVELTLOVGVGQ----- 828
 Db 14544 NSAGRS---APRESRPVKEQTMLPELDLR-GIYQKLVIAGADNIKVEIPVLGRP KP 1459
 QY 829 -LTEQRKD TLVRQ-----LAVLVN--LSDSIKVOKTRAHS---DLSTVIVFYV 871
 Db 14599 TVTWKGDQDIKTQTRVNFETTATSTILINECVSDSGPYPLTARNIVEGVDVITI 1465
 QY 872 QS--RPPFKVLKAAEVARNLHMELSKEDAFLLPKVL RVDTAGCLLKCSGHGCDPLTKR 929
 Db 14659 HDIPGPTTGPIKFEVS-----SDFVTFS----- 1468
 QY 930 CICSMLWMENLIQRYIWGDSEN---CEWSIFYVTVLAFTLIVLTGGFTWLICCCCKRQKR 986
 Db 14683 -----WDPPENDGGVPISNVVEMRQ-----TDSITTVELA-----T 1471
 QY 987 TKRKTKTYT-I LDNMDEQRMELRPKYGIKHSRSTHNSSLMWSES EF----- 1033
 Db 14715 TVTRTTYKATRLTGTLEYQPRVQAQNRYGVGPGIT---SAWIVANYPFKVPGPGT PQVT 1477
 QY 1034 --DSODQDTTSREKMERGNPKV---SNMGSI RNCAFSYSCK 1070
 Db 14772 AVTKDSMTISWHBEPLSDGGSPILGYHVRKERNGILMQTVSK 14813

RESULT 2
 S20301
 titin - rabbit (fragment)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: S20301; I46520
 R:Labelit, S.; Gautel, M.; Lakey, A.; Trinick, J.
 EMBL J. 11, 1711-1716, 1992
 A>Title: Towards a molecular understanding of titin.
 A:Reference number: S20897; MUID:92258380; PMID:1582406
 A:Accession: S20301
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-6805 <LAB>
 A:Cross-references: UNIPROT:Q28733; EMBL:X64696
 A>Note: The nucleotide sequence was submitted to the EMBL Data Library, February
 R:Labelit, S.; Barlow, D.P.; Gautel, M.; Gibson, T.; Holt, J.; Haiech, C.L.; Fran
 Nature 345, 273-276, 1990

A;Title: A regular pattern of two types of 100-residue motif in the sequence of titin.

A;Reference number: I46520; MUID:90238553; PMID:2129545

A;Accession: I46520

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 4235-5250 <LA2>

A;Cross-references: EMBL:X17329; NID:q1756; PIDN:CAA35207.1; PID:g930251

C;Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; pro

C;keywords: muscle

```
Query Match      3.7%; Score 208.5; DB 2; Length 6805;
Best Local Similarity 18.8%; Pred. No. 0.029;
Matches 212; Conservative 155; Mismatches 363; Indels 395; Gaps 59;

Qy 52 PPVVDCTAACCDLS-----SCDLA---WWFEGRCYLVSCHPKENCEP 90
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db FDISDIDADACSLSHIPLDGSGSNITNYIVEKCDVSRGDW-----VTALASVTKTCRI 915

Qy 91 KMGPIRSYLTFLVR-----PVORPAQLLDYGDMLNLRGSPGIWGDSPEDIR--- 138
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 916 GKLIPOGEYV-FRVAENRFGISEPLOSCKMAQF-----PFGV-PSEPKNAWTK 964

Qy 139 --KDLXFLGKDW-----GLEEMSEYXDDYRELEKDLLOPSGKQEPGSAEYTDMLLPGSE 192
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 965 VNKDCIFVANDRPDSDGSGSPITGYLIE-RKGRNSLLWVKANDTAVRSTYPCAGLVEGLE 1023

Qy 193 GAFNSSVGDSPVAPAEQTQDPDELHYLNESASTPAPKLPERSVLLPLPT-TPSSGEVLE-- 249
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1024 YSF-----RIVALNKAAGSSP-PSKPTTEYVARTPVDPGKPEVIDVT 1064

Qy 250 KEKASQLQEQSSNSGKEVL-----MPS-----HSLP-----PASLELSVVT 286
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1065 KSTVSLIWARPKHGGSKIIGFYVEACKLPGDKWVRNCTTTPHOIPHEEYVTGLEENAYQ 1124

Qy 287 VEKSPVLTVPSTGSEHSIPTPTSAAPSESPSELPIPTTAPRTVKELTVSAGDNLIIIT 346
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1125 QFRAIAKTAVNISQPSSELTPTVTHA--ENVPPRIDLS--VAMKSL--LTVKAGTNVCL- 1177

Qy 347 LPDNEVELKAFVAPAPPVETTYNYEWNLSHSPTDYQGEI---KQGHQTLNLSQSLVGLY 403
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1178 -----DATVFGKMPVTV-----SWK-----KEGVTLKPAEGIKWAMQRNLTLELF 1218

Qy 404 VFK-----VTVSSENAFGEFVNVTVKPARRVNLPPVAVV-----SPQLQE 444
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1219 SVNRKSGDYITITAEKSSGSKSATIKLVLDKVRPG-PPASVKINRWYSDRAMLSPEPLED 1277

Qy 445 LTLPLTSALIDGSDTDDTEIVSVHWEIEINGPFIEEKTSVDSPVLRLS---NLDPGNVSP 501
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1278 GGSEITNIVDKRETSS-----SRNNAQV-----SANVPTSCSVEKLIHGHEYQF 1322

Qy 502 RLTVTSDSGATNST-TAALI VNNAVDYPPVANAGPNHTITLPSNLSITLNGQSSDD---- 556
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1323 RICAEKNGYGVDPVFTPEPAIAKN--PYDPPGRCDDPPVLSNVTKDHTVSWKPPADGGSP 1380

Qy 557 -----HQVLVWLSLPGSGEGKHVVMQGVQTPYLHLASMQEG-DYTFOLKVYDTS 604
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1381 ITGYLLEKRETHAV---NWT---KVNRPKVIERTIKA-----TGLQEGTEYEFRTAINK 1429

Qy 605 S--RQQSTAXTVTVIQQENNRPPVAVAGPKDELIFPV-----ESATLDGSSSDHGH-- 654
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1430 AGPKPSDASKANVAQ-----DPLYPGPP---APPKVYDTRSSVLSWVKPKAYDGGSP 1481

Qy 655 IVFY-----HWEHVRGFSAVEMENIDKAIATVTGLQVGT-YHFRLTVKQOQGLSS 703
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1482 IIGYLVEKVRADTDNWRCNLP-----QKLQKTFEVTGLMENTEYQFVVAVNKGVSVD 1536

Qy 704 TS-----TLTVAVKKNNGSP----- 719
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1537 PSDVPDKHCPKDIILIPPEGLDADLRKTLILRAGVTNRLYVVPVGRPPPKITMSKPNVNL 1596

Qy 720 BARAG-----GRHVLIPLNN----- 734
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1597 RERIGLDIKSTDFDFTLRCEENVKYDAGKYLITLLENSCGKGGYTIIVKVLDPGPPVNV 1656
```

```
Qy 735 --SITLDGSRSTDDQRIYVSLWIRDGQSPAGDVIDGSD-----HSVALQLT 779
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1657 VKESIRDSAYITWDPPIV-----DQSGSPIINYVVEKRDADARKSWSTVTTTECSKTSFVS 1710

Qy 780 NLVEG-VYTFHLKRVTSQ-----ASDQTDATATVEVO-----PDPR 813
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1711 NLEBGSYFRRVVAENEYIGDPCGETDAVKASETPGPVVDLKVLTWKSSCNIGWKPKR 1770

Qy 814 KSGVELTTLQVGVQQLTEQRK-DTLVRQLAVLNLVDS-----IKVQKIRAHSDLS 864
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1771 SDGSRITGYV-VDFLTEENKQKRVKMSLSLOYSTKDLNEGKQYTFRVSANENGEGTPTS 1829

Qy 865 TVIVFYVQSPPPKVYLKAAEVAARNLHNR-----LSKEKADELL 902
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1830 EITV-----VAKDDVVPADLDLKDLPDLCLYLAKENSFRL 1864
```

RESULT 3

F90696

hypothetical protein EC0542 [imported] - Escherichia coli (strain O157:H7, substrain R1)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: F90696

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F90696

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-5291 <HAY>

A;Cross-references: UNIPROT:Q8X2T1; GB:BA000007; PIDN:BA033965.1; PID:g13360000; GSPDB:G

A;Experimental source: strain O157:H7, substrain R1MD 050952

C;Genetics:

A;Gene: EC0542

Query Match 3.7%; Score 208; DB 2; Length 5291;

Best Local Similarity 20.6%; Pred. No. 0.021;

Matches 173; Conservative 114; Mismatches 319; Indels 234; Gaps 34;

Qy 190 GSEGAFFNSSVGDSPVAPAEQTQDPDELHYLNESASTPAPKLPERSVLLPLPTPSSGEVLE 249

Dy :
Db 2854 GSDLAL-SGTSDDQPAITAIT---VTLNGQNYSATTDAS--GNWSVTVPASAVSALGEATY 2907

Qy 250 KEKASQLQEQSSNSGKEVLMPSHSLPSPASLELSSVTVEKSPVLTVPSTGSEHSIPTPT 309

Dy :
Db 2908 SVTASVTNAQNSSTASHNVQVNTALPGITIN---PVATDDIINASEAGSAQTISGQVT 2963

Qy 310 SAAPSESTPSELPISTPTAPRTVKELTVSAGDNLIIITLP-----DNEVELKAFVA-- 359

Dy :
Db 2964 GAAAGSVTVVGLGKTYTA-----TVQADLSNNVSVPAADWQALNGELTVNASVTNA 3016

Qy 360 -----PAPPVETTYNYE-WNLISHPTDYQGEIKQG----- 388

Dy :
Db 3017 VGNTSGSTRDITIDASLPGRLRVDTVAGDDVVNIIEHA---QAAQVITSGSGFAAGTALT 3073

Qy 389 --HKQTLNLSQSLVGLVVKVTVSSENAFGEFVNVTVKPARRVNLPPVAVVSPQLOBLT 446

Dy :
Db 3074 VINNQTVAAVTLANGSVSGVGPATDVSNWPAGTLNITVSGANSAG-TQTSITHP---LT 3128

Qy 447 LPLTSALIDGQSQSDTDD-----TETVSYHWE-----EING 475

Dy :
Db 3129 VDLTAVAI SMNSITSDDAINAAEKGAAALTLGSGTSGVGEAGQTVTVTGGKTYTTTVAAG 3188

Qy 476 PFTEEKTSVDSPLVRLNLPNGYSFRLTVDSDGATNSTTAA-----LIVNNAV 525

Dy :
Db 3189 SW-----STTVPAADLAALRDGDASQAVRVTVNNG--NSATATHEYSVDSAAPTVTINTI 3241

Qy 526 DYPPVANA-----GPNHTITLPSNLSITLNGQSSDDHQLVLY---EWSLPGSGEG 572

Dy :
Db 3242 ASDNIINASEAAAGVTVSGTSTAGTQGTQTLIVTLNGT-----NYQTTVTQDGSWSL----- 3291


```
QY 573 KHVVMQGVQTPYLHLSAMQEGDYYTFLQKVTDSRQOQSTAXVTVIYQPENRRPPVA---VA 629
Db 3292 -----TLPASDITLALANNCYTITATVSDLAGNLSGASKGVTV--DTTAPVIFSNIVA 3341
QY 630 GPD-----KELFFPVESATLDGSSSDDHGIVFYHWEHVRGVSASVEMENIDKAIATVTGL 684
Db 3342 GDDVINNVHEHIQAIIISGTATGAVAGDRLVVTIAGQVYVTTSDAGSNWSVGVPASVISGL 3401
QY 685 QVGTVHFRLTVKDOGLSSTST-----LTVAVKKNNSPPRARAGGRHVVLVLPN 733
Db 3402 ADGVTVISATITDSAGNSSTQTHNVQNTAAVLSVSTISGDNLLNAAEAG----- 3452
QY 734 NSITLDGSRSTDQRIVSYLWIRDCQSPAAGDVI-----DGDHSHVALQLTNLVEGVYTFH 789
Db 3453 SALTLSGT-----GTNFATGIVTVLLNGKYSATIO-----SNGSWSVN 3492
QY 790 LRVTSQOASGDDTATVATVEQDPDRKSLVELFLQVGVLQTEQRKDTLVRQLAVLLNV-- 847
Db 3493 VPAADVAALSDDGTSVTVSASAQD-----SAGNGNSSTQTHNVQNTAAVLSVST 3542
QY 848 LDSDKVKQIRAHSDLS-----TVIVFY-----VQSRPPFKV-LKAAEVA 886
Db 3543 TSGDNLLNAAEAGSALTLSGTCTNFATGTVTVTVLLNGKYSATIOQSGNSWSVNVPAADVA 3602

RESULT 4
S24600
projectin - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S24600
R.Fyberg, E.
submitted to the EMBL Data Library, May 1992
A:Reference number: S24600
A:Accession: S24600
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1742 <FVR>
A:Cross-references: UNIPROT:Q24463; EMBL:X66018; NID:g8377; PID:g8378
C:Genetics:
A:Gene: FlyBase:bt
A:Cross-references: FlyBase:FBgn0005666

Query Match 3.7%; Score 205; DB 2; Length 1742;
Best Local Similarity 19.6%; Pred. No. 0.0062;
Matches 187; Conservative 99; Mismatches 258; Indels 410; Gaps 48;

QY 80 VSCPHKENC-----EPKKG--PIRSYLTFLVLRVQRPQAQLLDYGDMMNLRGSPSGIWDG 132
Db 133 VSEVHKDGGCKLWKPKDGGEPVESYLVKEFDP-----DTGIW-- 171
QY 133 SPEDIRKDLXFLGKOWGLEEMSEYKDDYRELEKDLLQPSGKQEPGRGSARYTQWGLLPQSE 192
Db 172 -----LPVGRSD--GPEYNDGLVPGHD 192
QY 193 GAFN-----SSVGDSPAVPAET-----QQDPPELHYNESASTA-PKLPEPS-----V 234
Db 193 YKFRVAVKANKEGESE--PLETIGSLIAKDP-----FSVPTKGPVPEPTDWTNRKVEL 242
QY 235 LLPLPT-----TPSSGEVLE-KEKASQLQPSQSSSGKEVLMPSHSLPPASLESLSVTVEK 289
Db 243 ANPEPASDGGSPIQGYIVEKDKYSPLEWK-----ALETNSPT--- 280
QY 290 SPVLTV-----TPGSTHSIPTPTPSAAPSESTPSELPISTPTAPRTV 332
Db 281 -PTATVOGLIEGNEYQFRVVALNKGGLSEPSDPSKIFTAKPRY-----LAPKIDRRNL 332
QY 333 KELTVSAGDNLIITLPDNEVELKAFVAPAPPVETTYNNYEWNLISHPTDYQGEIKOGHKOT 392
Db 333 RNITLSSGTAL-----KLDANITGEPAPKV-----HLQSGKNVT 373
QY 393 LNLSQLSVGLYV-----FKVTVSSNAFGEFVNVTV--KPARRVNLPPVAVVSP 440
```

```
Db 374 IETPDYTKLVIRPQRSDSGEYLVATATNTSGKDSVLNVVVITDKPS-----PP-----NG 424
QY 441 QLQELTLPLTSALIDGQSQSTD--TEIVSYHWEI---NGPPIEEKTSVDSPLRLSLND 495
Db 425 PLQISDVHKGCHLWKAPSDGGGTPLEYFOIDKLEPETGCIWIPSCRSTE-PQVDVTGLS 483
QY 496 PGN-YSFRLTVTSDGATNSTTA--ALIVNNAVDPYPVANAGPNHTITLPQNSITLNGNQ 552
Db 484 PGNEYKFRVSAVNAEGESQPLVGDESIVARNPFDEP-----GKPN---LKATD 529
QY 553 SSDDH-----QIVLY-----EWSLG---PGSEGHVVMQGV--QTPY 584
Db 530 WDKDHVLDLAWTPPLIDGGSPISCVIIIEKQDKYKWERALDVPADQCKATIPDLVEGQTYK 589
QY 585 LHLSAMQEG----- 593
Db 590 FRVSAVNAAGTGPDSSTPPIIAKARNKPPIIDRSSILVEVRIKAGQSFPTDCKVSGEBAP 649
QY 594 -----DYTFOLKVYDSSSRQOS-----TAXVTV 615
Db 650 QTKWLLKKVEYKDNVKNVNTVNTKLKVNATSATSDSGIYTVFRENANGEDSADVKTV 709
QY 616 IVQENNRPPVAVAGPKELIFPVESATLDGSSSDDHGIVFYHWEHVRGVSASVEMENID 675
Db 710 IDKP---APP---NGPLKVDEINSECTLHWNPPDDGG-----QPID 746
QY 676 KAI-----ATVTGLQVG-----TYHRLTVKDOOGLSSTSLTVAV 711
Db 747 NYVVVKGLDETTGRWMTAGETDGPVTALKVGLLTPGHKYKFRVRAKNAQGTSEPLTTAAQI 806
QY 712 KKENNSPPRARAGGRHVLPNNISITLDGSRSTD--DQRIVSYLWIR---DQGPSAAGDV 766
Db 807 IAKN-----PFDPVTPKGTPIKDFKDFVDLEWTRPEADGGSPITGV 850
QY 767 IDGSDH-----SVALQLTN-----LVEGV-YTFHLRVTDGSGADTDTAT 805
Db 851 VEKRDKFSPOWEKCAEISDDITNAHVPLIEGLKYEFVRVAVNKAAGPGSPDAT 904

RESULT 5
B85547
probable RTX family exoprotein [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85547
R:Perina, N.F.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85547
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5188 <STO>
A:Cross-references: UNIPROT:Q8X4H5; GB:AE005174; NID:g12513368; PIDN:AAG54838.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z0615

Query Match 3.6%; Score 203.5; DB 2; Length 5188;
Best Local Similarity 20.8%; Pred. No. 0.036;
Matches 165; Conservative 109; Mismatches 296; Indels 223; Gaps 33;

QY 190 GSEGAFFSVGDSFAPVPAETQDDPELHYLINESASTPAPKLPEPSVLLPLPTPPSSGEVLE 249
Db 2854 GSDLAL-SGTSQDPAGTAIT---VTLNGQNTSATTDAS--GNWSVTVPASAVSALGEATY 2907
QY 250 KEKASQLQEQSSNSGKEVLMPSHSLPPASLESLSVTVEKSPVLTVTPGSTHSIPTPTPT 309
Db 2908 SVTASVNTQAQNSSTASINQVNTALPGIITN-----PVATDDIINASAGSAQTISGGQVT 2963
QY 310 SAAPSESTPSELPISTPTAPRTVKELTVSAGDNLIITLP-----DNEVELKAFVA-- 359
```


A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: C64483
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1009 <BUL>
A:Cross-references: UNIPROT:Q58863; GB:U67587; GB:L77117; NID:g1592103; PID:g1592109; TI
C:Genetics:
A:Map position: FOR1437031-1440060

```
Query Match          3.6%; Score 200.5; DB 2; Length 1009;
Best Local Similarity 22.6%; Pred. No. 0.0049;
Matches 127; Conservative 70; Mismatches 22; Indels 137; Gaps 27;

Qy 298 GTSHTSPTPTSAAPSESTSELPSTPTAPRTVKELTVSAGDNLITLPDNEVELKAF 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 GNEVGIP-PPGVGGSSYP-QIVASPTIAP-TYYPINIT-----I 540

Qy 358 VAPAPPVETTYNENLISHPTDQGE-----IK-QGHKQTLNLSQSLVGLVFKVTYSS 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 VEP-----SWTWYHHI-----YFGDGGVMIKPKSPYTFTHYTSSEGYPIYMKYIT 590

Qy 412 ENAFEGEVNVTVKPARVNLPVAVVSPQLQELTL-----PUTSALIDGSGSTDDTE---- 464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 591 AENMKTVVINDKPIAKLIYP-----NPASYKOTISFSPINSYDPPDANRMIPEYNYGT 646

Qy 465 IVSYHWEIEINGPFIE-----EKTSVDS-----PVLRLSNLDP-GNYSFRLTVTDS 508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 647 LIGYVPISPNPMARIYGNLTVDYDSNGVAVWYSSNELTIISKSPFIGNYAKLAVWDG 706

Qy 509 DGATNSTAALIVNAVVDYPPVA-----NAGPNHTITLPQNSITLNGNSSDDHQIIVL 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 707 MGVNSTTVKFSV---INRPVAFIYYPDPKPEPNEDEVFVSQSYDPEG-----ETAY 756

Qy 562 YEWSLGPGSECKHVVMQGVQPYLHLSAMQSGDYTFQLKVTDSRQSQSTAXTVTVIQQEN 621
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 757 YIMFNGDT-----VINTTDT-IYHKKYRGYTVTLTVFDKYNASSISKQIVVGG-- 808

Qy 622 NRPPVAVAGPKELI-----FPVESATLDGSSSDDHGIYFYHWEHVRGSPSAVEMENID 675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 809 -----ILANFYEIINENTVRFEDTSVAPGK-----IISHWDFGSGSTSEQNPIH 856

Qy 676 KAIATVGLQVTHFRITLVKQDQGLSSTSTLTVAVKKENNSPPRARAGGRHVLVLPNS 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 857 -----TYSKEGAFVTLTVKSDTNLTDSVTKIIVISPPPKYPPVADFTYK---VINGTT 907

Qy 736 ITLDSRSTD-QQRIVSVLWIRDOQSPAAGDVIDGSDHSVALQLTNLV-----EGVYT 787
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 908 VEFNANLSYDEGGYIKYIY-----DFGDTPTANTTNPITIIHKKYKKEGVYP 954

Qy 788 FHLRVDSQSGASDSDTATVEV 808
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 955 VTLTVINDNLSDSRKLIINI 975
```

RESULT 10
A46405
C:protein alpha antigen - *Streptococcus* sp. (strain A909, group B, type Ic)
C:Species: *Streptococcus* sp.
C:Date: 21-Sep-1993 #sequence revision 20-Sep-1999 #text_change 03-Nov-2000
C:Accession: A46405; B46405; T30854
R:Michel, J.L.; Mañoff, L.C.; Olson, K.; Kling, D.E.; Kasper, D.L.; Ausubel, F.M.
Proc. Natl. Acad. Sci. U.S.A. 89, 10060-10064, 1992
A:Title: Large, identical, tandem repeating units in the C protein alpha antigen gene, B
A:Reference number: A46405; MUID:93066179; PMID:1438195
A:Accession: A46405
A:Molecule type: DNA
A:Residues: 1-1020 <MIC>
A:Cross-references: GB:M97256; NID:g153558; PIDN:AAA26848.1; PID:g153559
A:Note: sequence extracted from NCBI backbone (NCBIN:117590, NCBIN:117592, NCBIN:117593)
A:Note: the source is designated as *Streptococcus agalactiae*, group B, strain A909; ATCC
C:Genetics:

A:Gene: bca
C:Keywords: tandem repeat
F:227-308/Region: 82-residue tandem repeat
F:309-390/Region: 82-residue tandem repeat
F:391-472/Region: 82-residue tandem repeat
F:473-554/Region: 82-residue tandem repeat
F:555-636/Region: 82-residue tandem repeat
F:637-718/Region: 82-residue tandem repeat
F:719-800/Region: 82-residue tandem repeat
F:801-882/Region: 82-residue tandem repeat
F:883-964/Region: 82-residue tandem repeat
F:965-975/Region: 11-residue partial tandem repeat
F:997-1014/Domain: transmembrane #status predicted <TRM>

Query Match 3.5%; Score 195.5; DB 2; Length 1020;
Best Local Similarity 23.0%; Pred. No. 0.0093;
Matches 173; Conservative 85; Mismatches 293; Indels 201; Gaps 41;

```
Qy 164 EKDLLQPSGKQE--PRGS-----AEYTDGMLLPGSEGFNSVSGDSP--AVPAETQQDPPEL 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 DKDKYDPTGETTVPGQTPVSDKEITDLVKIPDGSKGVPTVVGDRPDTNVPDGHKVTVEV 288

Qy 216 HYLNESAS-----TPAPKLPERSVLLPL--PTTPSSGEVLEKEKASQLEQSSNSG 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 TYPDGTKDTVEVTVHVTPKP-VPDKDKYDPTGGETTVPQTPVSDKEITDLVKIPDGSKG 347

Qy 266 KEVLM-----PSHSUPPASLELSSTVVEKSPVLTVTPGS-----TEHSIPTPTTSAAPSES 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 VPTVVGDRPDTNVPD-----HKVTVE---VTYPDGKTDTVEVTVHVTPKPVPDKDKYD 399

Qy 317 TPSELPISPTAPTAKELTVSA-----GNNLI---ITLDP-- 349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 TGGETTVPQTPVSDKEITDLVKIPDGSKGVPTVVGDRPDTNVPDGHKVTVEVTPDGT 458

Qy 350 -NEVELKAFVAPAP-PVETTYNENLISHPTDYQGEIKQG-----HQQLMLSQL---SV 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 KDTVEVTVHVTPEVPDKKYD-----PTGETTVPGQTPVSDKEITDLVKIPDGSK 510

Qy 401 GLYV-----FKVTVSSENAFG-----EGFNVTVKPARVNLPVAVVSPQL 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 511 GVPVVGDRPDTNVPDGHKVTVEVTPDGTKDTVEVTVHVTPKP-----VPDKDKYDPTG 565

Qy 443 QELTLPITLSALIDGSGSTDDTEIYSHWEIEINGPFIIEKTSVDSQSVLRLSNLDGNYSF 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 GETTVPGQTPVSD-KEITDLVKIP-----DGSKGVPTVVGDRPDTNVP-----PDGHKVT 613

Qy 503 LTVTSDGATNSTTAALI-----VNAVDPVPPVANAGPNHTITLPQNSITLNGNQSDSD 556
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 614 VEVTYPDG-TKDTVEVTVHVTPEVPDKKYDPTGG-----ETTVPG-----GTFVSDK 661

Qy 557 HQIVLYEWSLPGSGEGKHVVM---QGVQTPYVHLSAMQSGDYTFQLKVTDSRQSQSTAXV 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 EITDLVK--IPDGSKGVPTVVGDRPDTNVP-----GDHKVTVEVTVPDGKTDTVEV 710

Qy 614 TVIVQPENRPPVAVAGPDKELIAPPV--ESATLDGSSSDDHGIYFYHWEHVRGSPSAVEM 671
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 711 TVHVTPEK-----PV-----PDKDKYDPTGGETTVPQTPVSDK-----EITDLVKI 751

Qy 672 ENIDKATATVGLQ-----VGTVHFRITLVKQDQGLSSTSTLTVAV-----KKNNSPPR 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 752 PDGSKGVPTVVGDRPDTNVPDGHKVTVEVTVPDGKTDTVEVTVHVTPEVPDKDKYD-- 809

Qy 721 ARAGGRHVLVLPNNSITLDGSRSTDDQRIYSVLWIRDOQSPAAGDVIDGSDHSVALQLTN 780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 810 --TGG-----ETTVPGQTPVSDKEITDLVKIPDGSKGVPTVVGDRPD-----TN 851

Qy 781 LVEGVYTFHLRVTDGSGASDSDTATVEVQDP 812
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 852 -VPDGHKVTVEVTVPDGKTDTVEVTVHVTPEK 882
```

RESULT 11
T34434

Db 810 ITKDSVTLHWDLPLIDGSGSITNIVIVEKREATRKSYSVTTKCHKCYKVTGLSEGCEYF 869
Qy 691 FRLTVKDDQOGLSSTSTLTVAVKKENNSPPRARAGRHVLVLPNNISITLDGSRSDDDQRIV 750
Db 870 FRVMAENEYIGIBSETPKEPKASEAFSP-----PDSLNIMDTIKST-----V 912
Qy 751 SYLWIR---DQOSPAGDVID-----GSD---HVALQ-----LTNLVEG-VYTFHLRVTD 794
Db 913 SLAMPKPKHGGSKITGYVIEAQRKGSQDQWTHITTVKGLBCVVRNLTEGEYTFQVMAVN 972
Qy 795 SQGAS 799
Db 973 SAGRS 977

RESULT 13
E90696
hypothetical protein ECs0541 [imported] - Escherichia coli (strain O157:H7, substrain R1
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: E90696
R;Hayaashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E90696
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1461 <HAY>
A;Cross-references: UNIPROT:Q8XD19; GB:BA000007; PIDN:BA033964.1; PID:g13359999; GSPDB:G
A;Experimental source: strain O157:H7, substrain R1MD 050952
C;Genetics:
A;Gene: ECs0541

Query Match 3.4%; Score 190.5; DB 2; Length 1461;
Best Local Similarity 21.1%; Pred. No. 0.029;
Matches 139; Conservative 81; Mismatches 214; Indels 225; Gaps 28;

Qy 191 SEG--AFNSSVGDSPPAVPAETQODPELHLYNESASTPAPKLPERSVLLPLPTTSSGEVL 248
Db 442 AEGYAVNVSVDRAQ-----NTTSHSANFTVDTSPAPVSVNTV-----AGDDIL 486
Qy 249 EKEKASQLQOQSSNSGKE-----VLMPSHSLPPASLELSSTVTEKSPVLTVTPGSTE 302
Db 487 NNAEQVAQAIIISGVSGASPGDVTTVKLGTHVLGIVLADGSMNVALDPAVTRTLDRGAN 546
Qy 303 SIPTPTSAAPSESTPSE-----LPISPTTAPRTVKELTVSAGDNLITLDPNEVELKAFV 358
Db 547 TIFVTVTDAAGNTGAASRAITLVGVSPLIITNTV-----SGDDII-----SGA 589
Qy 359 APAPPVETTYNWEWLISHPTDYQGEIKQGHKQTLNLSQSLVGLYVFKVTVSSENAFGE 418
Db 590 EKGAPLTLTGSTQ-----QAEQTGTVTVTLAGOS-----FTTIVQADGSM--- 629
Qy 419 FVNVTVPARRVNLPPVAVVSPOQLTLPLTSALIDGSOSTDDTEIVSVHWEINGPFI 478
Db 630 --SLTVPAAANGNLPCGAVA-----ITASVTLDSGNTGNT----- 662
Qy 479 EEKTSVDSPLRLNLNLPNGYSFRLTVTDSGATNSTTAALIVNNA---VDYP--PVANA 533
Db 663 SRTITVDSQAPALS-IDP-----LTADNIIAASGGDLPTITGTTDA 703
Qy 534 GPNHTITLPONSITLNGQSSDDHQIVLYEWSLPGSGEGKHVMQGVQTPVLHLHSAQEG 593
Db 704 QPGQGTVTTLNGQTYQGVQPDG-----TWS-----VTVPAANGALADG 743
Qy 594 DYTFLKVTDSRRQQSTAXVTIVIQENNRPPVAVAGP---DKELIFP-----VESATL 644
Db 744 NATVTASVNDVAGNPSVSRVALV---DATPPVVTINPVDNVTINTPEHAQAQIIISGTV 800
Qy 645 DGSSSSDDHGIVFYHWEHVRGSPSAVEMENIDKA-----IATVTGLQVGTYH 690

Db 801 TGAQAGD---IV-----TVTLLNNVDYTTVVDGSGNWSLGVSPASVVSGLADGSYP 846
Qy 691 FRLTVKDDQOGLSSTSTLTVAVKKENNSPPRARAGRHVLVLPNNISITLDGSRSDDDQRIV 750
Db 847 VSVSVTDKAGNTGQSLSLTVTV---NTAAP-----LIGINSI----- 879
Qy 751 SYLWIRDDQSPAGDVIDGSDHVALQTLNLVEGVYTFHLRVTDSCASDSDTDTATVEVQ 809
Db 880 -----AGDDVINASEKGADLQITG-----TSDQPVNTAITVTLNGQ 915

RESULT 14
A85547
hypothetical protein Z0609 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: A85547
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A85547
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1461 <STO>
A;Cross-references: UNIPROT:Q8XD19; GB:AB005174; NID:g12513364; PIDN:AAG54837.1; GSPDB:B
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z0609

Query Match 3.4%; Score 190.5; DB 2; Length 1461;
Best Local Similarity 21.1%; Pred. No. 0.029;
Matches 139; Conservative 81; Mismatches 214; Indels 225; Gaps 28;

Qy 191 SEG--AFNSSVGDSPPAVPAETQODPELHLYNESASTPAPKLPERSVLLPLPTTSSGEVL 248
Db 442 AEGYAVNVSVDRAQ-----NTTSHSANFTVDTSPAPVSVNTV-----AGDDIL 486
Qy 249 EKEKASQLQOQSSNSGKE-----VLMPSHSLPPASLELSSTVTEKSPVLTVTPGSTE 302
Db 487 NNAEQVAQAIIISGVSGASPGDVTTVKLGTHVLGIVLADGSMNVALDPAVTRTLDRGAN 546
Qy 303 SIPTPTSAAPSESTPSE-----LPISPTTAPRTVKELTVSAGDNLITLDPNEVELKAFV 358
Db 547 TIFVTVTDAAGNTGAASRAITLVGVSPLIITNTV-----SGDDII-----SGA 589
Qy 359 APAPPVETTYNWEWLISHPTDYQGEIKQGHKQTLNLSQSLVGLYVFKVTVSSENAFGE 418
Db 590 EKGAPLTLTGSTQ-----QAEQTGTVTVTLAGOS-----FTTIVQADGSM--- 629
Qy 419 FVNVTVPARRVNLPPVAVVSPOQLTLPLTSALIDGSOSTDDTEIVSVHWEINGPFI 478
Db 630 --SLTVPAAANGNLPCGAVA-----ITASVTLDSGNTGNT----- 662
Qy 479 EEKTSVDSPLRLNLNLPNGYSFRLTVTDSGATNSTTAALIVNNA---VDYP--PVANA 533
Db 663 SRTITVDSQAPALS-IDP-----LTADNIIAASGGDLPTITGTTDA 703
Qy 534 GPNHTITLPONSITLNGQSSDDHQIVLYEWSLPGSGEGKHVMQGVQTPVLHLHSAQEG 593
Db 704 QPGQGTVTTLNGQTYQGVQPDG-----TWS-----VTVPAANGALADG 743
Qy 594 DYTFLKVTDSRRQQSTAXVTIVIQENNRPPVAVAGP---DKELIFP-----VESATL 644
Db 744 NATVTASVNDVAGNPSVSRVALV---DATPPVVTINPVDNVTINTPEHAQAQIIISGTV 800
Qy 645 DGSSSSDDHGIVFYHWEHVRGSPSAVEMENIDKA-----IATVTGLQVGTYH 690
Db 801 TGAQAGD---IV-----TVTLLNNVDYTTVVDGSGNWSLGVSPASVVSGLADGSYP 846
Qy 691 FRLTVKDDQOGLSSTSTLTVAVKKENNSPPRARAGRHVLVLPNNISITLDGSRSDDDQRIV 750

```

Db      847 VSVSVTDKAGNTGQSGLTVTV---NTAAP-----LIGINSI----- 879
Qy      751 SYLWTRDQGSPPAGVDVDSHVSVALQLTNLVEGVYTHLRYVDSQASDTRDTATVEVQ 809
Db      880 -----AGDDVINASEKAGDLQITG-----TSDQPVNTAITVTINGQ 915

RESULT 15
AH2493
hypothetical protein all7128 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2493
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
  Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
  DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
  A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2493
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3083 <KUR>
A:Cross-references: UNIPROT:O8YL10; GB:BA000020; PIDN:BAW78212.1; PID:gi17135666; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all7128
A:Genome: plasmid

Query Match      3.4%; Score 190; DB 2; Length 3083;
Best Local Similarity 21.1%; Pred. No. 0.091;
Matches 167; Conservative 96; Mismatches 302; Indels 226; Gaps 35;

Qy      148 WGLEEMSEYXD--DYRELEKDLQPSGKQPRGSAEYTDWG----- 186
Db      1976 YGSLALSSYVDIGVRYNSL--DL-----ATSGEISTWGPVSHVRYNIERFDITGTR 2025
Qy      187 -----LLPGSEGAFNSSVG---DSPAVPAET-----OODPE-----LHYLE 220
Db      2026 YGDELLGGNLDKLTGGGGADTLKAGLGNDDTVILAAQTAGGSKIEDDGDNDTLDLTINL 2085
Qy      221 SASTPA-----PKLPERSVLLPLP-----TTPSSGEVL-----EKEKASQLQEQSSNSSG 265
Db      2086 SLSTFTTGTAGTQRL-GYTLILDLDNQDGIITTPESDLSIINFNSSAGTGFIKVDNLSG 2144
Qy      266 KEVLMPSHSLPPASLESLSVTVKSPVLVTPGSTEHSIPTPTSAAPSESTPSELPISP 325
Db      2145 TDIL---NKLFGNSANQAPVT-QANKVLTV---AEDSVTTPLAJATPTDNDLLTITI 2196
Qy      326 TTAPRTVKELTVSAGDNLLIITLPDNEVELKAFVAPAPPVETTYNYEWNLI SHPTDYQGEI 385
Db      2197 TAVPEASKG-----IIRLPDNTV-----VTVN----- 2218
Qy      386 KQCHKQTLNLSQLSVGLYVFKVTVSSENAFGE-GFVNVTVPKARKVNLPPVAVVSPOLQE 444
Db      2219 -----TLLTQQLT-----SLVFSVVNANGSAGSFYTVSDGK-----GTASQT 2259
Qy      445 LTLPLTSA-----LIDGSQSTDDT-----EIVSY----- 468
Db      2260 ITLEITAVNDAPTLANAJANQATATEDTAFTTIPANTFTDVGADGALTVSATLADGANLP 2319
Qy      469 HWEENGPFEKTSVDSVPLRLSNLDPCGNSFRLTVTDSGATNSTTAALIVNNA--VD 526
Db      2320 NWLSFNP---STRFTIGTP-----TNNSVGTVNIRVATDNAGASVDVFTLTVANSDDTND 2372
Qy      527 YPPVANAGPNH-----TITLPONSITLNGQSDDDHQIVLYEWSLGPSEGKHVVNQ 578
Db      2373 APTLENAIANQATATEDSAFTTIPANTFA-----DVDAGDTLTYSATLADGADLLNWLNF 2427
Qy      579 GYQTPYVHLHSAMQEGDYTFOLKV--TDSRQOSTAXVTIVOPENNRPPVA-----VAG 630
Db      2428 NPSTRFTSCGTPNDVEGTINIKVKTATDNAGASLIDFTLTINTNDAPTAVANAIANQATAT 2487
```

```

Qy      631 PDKELIFPVESATLDGSSSDDHGIVFVWHEVHVGPSAVEMENIDKAIATVTGL----QV 686
Db      2488 EDTAFNQIPA---DAFNDVDVTGDTLTYTATLENGDELPSWLTFDAATRTTSGTPTNSEV 2544
Qy      687 GTYHRLTVKDDQOQGLSSTSTLTVAVKKNNSP-----PRRAGGGRHVLVLPNNSITL 738
Db      2545 DTLSIKVIATDKSQASASNVFTLTVLNTNDAPTLENAIADQTATEDSTSFIIPVNTEA- 2603
Qy      739 DGSRSSTDQRIVSY-LMIRDQSPAGDVIDGSDHSVALQLTNLVEGVYTHLRYVTDQSG 797
Db      2604 ----DVDADDILAYSATLEBGAALPSWLTFNPTNRTTFAGTPTINSEVGTLNLIKVIATDKSS 2659
Qy      798 ASDTDTATVEV 808
Db      2660 ANVSDVFTLTV 2670
```

Search completed: October 12, 2005, 10:22:42
Job time : 58 secs